Deltamodulation mit adaptivem PARCOR-Prädiktionsfilter

DISSERTATION

zur Erlangung des Grades eines
Doktor-Ingenieurs
der Abteilung für Elektrotechnik
an der Ruhr-Universität Bochum

von

BERND SCHLEIFENBAUM

Schleswig

Bochum 1981

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41549615_0027

Deltamodulation mit adaptivem PARCOR-Prädiktionsfilter

DISSERTATION

zur Erlangung des Grades eines
Doktor-Ingenieurs
der Abteilung für Elektrotechnik
an der Ruhr-Universität Bochum

von

BERND SCHLEIFENBAUM

Schleswig

Bochum 1981

UNIV.
LEIDEN
BIBL.

Dissertation eingereich am: 6. Januar 1981

Tag der mündlichen Prüfung: 28. Januar 1981

Referent: Prof. Dr. sc. techn. A. Fettweis

Korreferent: Prof. Dr.-Ing. J. Blauert

Inhaltsverzeichnis

		Seite
1.	Einleitung	7
1.1	Überblick über die Entwicklung der DM	8
1.2	Ziel dieser Arbeit	9
2.	Ergebnisse der Rate Distortion Theorie (RDT)	12
2.1	Charakterisierung von zeitdiskreten, stationären, gaußschen Signalquellen mit Gedächtnis	14
2.2	Definition der Rate Distortion Funktion (RDF) und Variationsproblem zur Berechnung der RDF für eine Signalquelle ohne Gedächtnis	15
2.3	R (D) für stationäre, gedächtnisfreie, gaußsche Signalquellen mit dem Verzerrungsmaß mittlerer quadratischer Fehler (MQF)	17
2.4	R(D) für gaußsche Signalquellen mit Gedächtnis	19
2.5	R(D) für überabgetastete Signalquellen bei vorgegebener Kanalkapazität C	23
2.6	Autoregressive Signalquellen	25
2.6.1	Definition und RDF von gaußschen, autoregressiven Signalquellen	25
2.6.2	Baumcodierung von gaußschen, autoregressiven Signalquellen	27
3.	Beschreibung des Sprachprozesses	33
3.1	Autoregressive Signalquelle als Modellsignalquelle für den Sprachprozeß	34
3.2	Koeffizientenbestimmung der Modellsignalquelle aus realem Sprachsignal	35
3.3	PARCOR-Analysefilter zur rekursiven Lösung des linearen Gleichungssystems der Autokorrelationsmethode	37
3.4	PARCOR-Synthesefilter	40
4.	DM mit adaptivem Rekonstruktionsnetzwerk	43
4.1	Vergleich von DM und optimaler Baumcodierung	46
4.2	DM bei niedrigen Abtastraten	49
4.3	Struktur des adaptiven Rekonstruktionsnetzwerkes	51
4.4	Vereinfachte und schnelle Verfahren zur Bestimmung der Reflexionskoeffizienten	53

		Seite
5.	Stabilität des Deltamodulationsprozesses	57
5,1	Grenzzyklusschwingungen (Grenzzyklen) in nichtlinearen Abtastsystemen	57
5,2	Existenzbedingung für beliebig vorgebbare Moden bei einem Deltamodulator mit M-stufigem PARCOR-Synthesefilter und adaptiver Quantisierungsstufenbestimmung bei Eingangssignal $x(t)\equiv 0$	59
5,3	Auswahl der zu testenden Grenzzyklusmoden	62
5,4	Berechnung von Stabilitätsdiagrammen für einen Deltamodulator mit M-stufigem PARCOR-Synthesefilter	64
5,4,1	Stabilitätsdiagramm für 2-stufiges PARCOR-Synthesefilter	64
5,4,2	Ermittlung des Reflexionskoeffizientenbereichs bis M=4 bei der PARCOR-Analyse von Sprachsignalen (Abtastrate = 16 kHz und Tiefpaßfilterung bis 3,4 kHz)	66
5,4,3	Untersuchung einiger ausgewählter unsymmetrischer Moden	68
5,4,4	Systematische Untersuchung aller symmetrischen Moden bis zur Periodenlänge L=12 bei 2-stufigem PARCOR-Synthesefilter und fester Quantisierungsstufe (g_n=1)	70
5,4,5	Untersuchung der Änderung der Stabilitätsbereiche für einige ausgewählte symmetrische Moden bei 4-stufigem PARCOR-Synthesefilter mit fester Quantisierungsstufe (g_n=1)	74
5,4,6	Änderung der Stabilitätsbereiche einiger ausgewählter Moden bei 2- und 4-stufigem PARCOR-Synthesefilter und adaptiver Quantisierungsstufeneinstellung	77
5,5	Existenzbedingung für beliebig vorgebbare Moden bei einem Deltamodulator mit M-stufigem PARCOR-Synthesefilter und adaptiver Quantisierungsstufeneinstellung bei $x(t)=\text{const.}\neq 0$	83
5,6	Anregung von Grenzzyklen	85

		Seite
6.	Optimierung und Beurteilung der Rekonstruktionsgüte von synthetischen Sprachsignalen	87
6.1	Gaußsches, autoregressives Testsignal	87
6.1.1	Synthetischer Vokal |ä|	89
6.1.2	Synthetischer stimmloser Laut |sch|	92
6.2	Frequenzgewichtete Gütekriterien	95
6.3	Autoregressives Testsignal mit impulsförmiger Anregung	99
7.	Festkommasimulation eines Deltamodulators mit adaptivem PARCOR-Synthesefilter und Augenblicksadaption der Quantisierungsstufe	103
7.1	Simulation einer PARCOR-Filterstufe	103
7.2	Stabilität von Grenzzyklen in Abhängigkeit der Wortlänge der Reflexionskoeffizienten	105
7.3	Beschränkung des Wertebereiches von k_1 und Änderung der Stufenzahl des PARCOR-Synthesefilters in Abhängigkeit von k_1	109
7.4	Einfluß von Kanalstörungen	115
8.	Ermittlung der subjektiv empfundenen Sprachqualität	117
8.1	Ablauf und Ergebnis des Hörversuchs	119
9.	Zusammenfassung	124
	Literaturverzeichnis	127
	Anhang A1 - A8	132

1. Einleitung

Durch die wachsende Bedeutung der digitalen Nachrichtenübertragung wächst auch das Interesse an der digitalen Verschlüsselung von kontinuierlichen Signalen. Das Gebiet der Nachrichtentechnik , das sich damit beschäftigt, ist die Quellcodierung. Das Ziel der Quellcodierung ist es, die in der Signalquelle vorhandene Information bei möglichst großer Übertragungsqualität mit möglichst wenigen digitalen Zeichen zu verschlüsseln. Bekannte Quellcodierungsverfahren, die den zeitlichen Verlauf von kontinuierlichen Signalen codieren, sind:

Pulscodemodulation - PCM,

Differentielle Pulscodemodulation - DPCM,

Adaptive differentielle Pulscodemodulation - ADPCM,

Lineare Deltamodulation - LDM,

Adaptive Deltamodualtion - ADM.

Diese Verfahren unterscheiden sich hinsichtlich der Abtastrate und der Anpassung des Codierervorganges an gewisse statistische Eigenschaften der Signalquelle. Als Signalquelle wird in dieser Arbeit hauptsächlich die menschliche Sprache betrachtet, und zwar für einen Frequenzbereich bis 3.4 kHz, so wie er bei Fernsprechkanälen zugrunde gelegt wird.

Die Verfahren der Pulscodemodulation arbeiten mit einer Abtatsrate, die der doppelten Signalbandbreite entspricht, während die Deltamodulationsverfahren üblicherweise bei höheren Abtastraten arbeiten. Die grundsätzlichen Vorteile der DM gegenüber PCM-Verfahren sind:

1. Es wird statt eines aufwendigen Quantisierers nur ein einfacher Komparator zur Analog/Digital(A/D)-Wandlung benötigt (1-bit-Quantisierer).

2. Es ist deswegen kein Abtasthalteglied notwendig.

3. Da zu jedem Abtastzeitpunkt nur ein bit an Information übertragen wird, ist keine Wortsynchronisation nötig.

4. Durch die höhere Abtastrate kann das Bandbegrenzungs- bzw. Glättungsfilter einfacher ausgeführt werden.

Der wesentliche Nachteil der DM ist die hohe Abtastrate, durch die nicht die optimal mögliche Rekonstruktionsgüte bei einer bestimmten vorgegebenen Kanalkapazität erreicht werden kann. Eine weitere Einschränkung, die die DM aber mit den DPCM- und ADPCM-Verfahren gemeinsam hat, sind gewisse Anforderungen an das Signalspektrum. Diese Anforderungen lassen aber die Datenreduktion gegenüber der PCM erst zu.

1.1 Überblick über die Entwicklung der DM

Die DM wurde 1946 in Frankreich patentiert [1]. Eine erste grundlegende Untersuchung führte de Jager [2] 1952 durch. In dieser Arbeit wurden schon verschiedene Rekonstruktionsnetzwerke betrachtet. Der 1-bit-Quantisierer war allerdings fest, so daß diese ersten Deltamodulationssysteme nur eine ungenügende Dynamik besaßen. Dieser Nachteil wurde durch die Einführung eines adaptiven Quantisierers durch Winkler [3] 1963 behoben. Die Adaption der Quantisierungsstufe wurde in diesem System durch die Auswertung der letzten gesendeten bits durchgeführt. In den folgenden Jahren erschienen eine Reihe von Arbeiten [4-10], die auf ähnlichen Prinzipien beruhten. Zusätzlich zu den Versuchen, die Qualität des rekonstruierten Signals zu verbessern, erschienen Untersuchungen über die Art der Quantisierungsverzerrungen und ihre Berechnung in Deltamodulationssystemem [11-15]. Aufgrund der komplizierten mathematischen Beschreibung eines adaptiven Deltamodulators ist eine analytische Optimierung der Systemparameter nicht möglich, besonders dann, wenn man die Codierung eines Zufallssignals betrachtet. Die Berechnung des auftretenden Quantisierungsgeräusches erfolgt daher nur unter

bestimmten Annahmen über den Zufallscharakter des Eingangssignals und die Art des Codierungsprozesses.

Eine Ausnahme bildet in diesem Zusammenhang die Arbeit von Slepian |15|. In dieser Arbeit wird für das einfachste mögliche Deltamodulationssystem - fester Quantisierer mit idealem Integrator als Rekonstruktionsnetzwerk - die Verbundwahrscheinlichkeitsverteilung von Eingangssignal und Ausgangssignal berechnet. Als Signalquelle wurde gaußsches Zufallssignal mit rationalem Leistungsspektrum betrachtet. In diesem Fall ist es also möglich, den einzigen frei wählbaren Parameter (abgesehen von der Abtastrate) - die Quantisierungsstufe - exakt zu optimieren. Für einen nichtidealen Integrator als Rekonstruktionsnetzwerk konnte allerdings keine Lösung angegeben werden.

Eine weitere, nicht auf die DM beschränkte Schwierigkeit bei der Optimierung besteht in der Wahl des Gütekriteriums. Für hohe Abtastraten (>40 kHz) stellte sich aber doch heraus |16-18|, daß auch mit aufwendigeren Deltamodulatoren eine bestimmte Qualitätsgrenze nicht überschritten werden kann.

Es gibt nämlich eine feste, nicht zu überschreitende Qualitätsgrenze für die Signalrekonstruktion, wenn man eine bestimmte Signalklasse betrachtet und zur Übertragung nur eine bestimmte Kanalkapazität zur Verfügung steht. Diese Grenze läßt sich mit Hilfe der Rate Distortion Theorie |19| berechnen. Es läßt sich ebenfalls zeigen |16,17|, daß man die zur Verfügung stehende Kanalkapazität am besten ausnutzt, wenn man das Signal mit der niedrigst möglichen Abtastrate (Nyquistrate) abtastet. Wenn man dann, wie bei der DM, nur ein bit pro Abtastwert überträgt, kann man mit den bekannten Deltamodulationssystemen keine befriedigenden Rekonstruktionsgüten erreichen.

1.2 Ziel dieser Arbeit

In dieser Arbeit wird versucht, den Deltamodulationsprozeß für Sprachsignale im Bereich niedriger Übertragungsbitraten (8-20 kbit/s) zu optimieren. Da eine analytische mathematische Beschreibung dieses Prozesses für Zufallssignale (Sprache) nicht möglich erscheint, wird eine indirekte Optimierung durchgeführt. Grundlage dieser Optimierung sind Ergebnisse aus der Rate Distortion Theorie und der Quellcodierung für zeitdiskrete und am-

plitudenkontinuierliche Zufallsprozesse. Dabei werden bestimmte Modellvorstellungen dieser Prozesse benötigt.

Zur Beschreibung der Erzeugung menschlicher Sprache gibt es gute Modellvorstellungen |20,21|. Gerade diese Modelle sind aber auch in der Rate Distortion Theorie und der Quellcodierung von großer Bedeutung.

Durch die Betrachtung der absoluten Qualitätsschranke, die durch die Rate Distortion Theorie berechnet werden kann, ist es dann sogar möglich, nicht nur eine Parameteroptimierung für ein bestimmtes Codierungssystem vorzunehmen, sondern die Optimierung auch durch diese absolute Qualitätsschranke zu überprüfen. Diese Aussagen gelten aber nur für ideale Modelle und bestimmte objektive Gütekriterien.

Eine subjektive Beurteilung der Rekonstruktionsgüte ist sehr viel schwieriger und erfordert umfangreiche Testreihen. Eine allgemein verbindliche Vorschrift für solche subjektiven Testreihen existiert allerdings noch nicht, so daß eine solche Untersuchung keine allgemein gültige Aussagekraft besitzt. Trotzdem werden in dieser Arbeit auch beschränkte subjektive Hörtests durchgeführt. Dabei soll ein Vergleich mit zwei anderen bekannten Quellcodierungsverfahren (Krümmungsadaptive DM, ADPCM) durchgeführt werden. Zur Objektivierung dieser Testergebnisse werden darüber hinaus an das Sprachsignal und das untersuchte Quellcodierungsverfahren angepaßte objektive Gütekriterien betrachtet.

Auf dem in dieser Arbeit untersuchten Gebiet der Codierung des zeitlichen Verlaufes von Sprache bei niedrigen Übertragungsbitraten sind bisher schon eine Reihe von Untersuchungen durchgeführt worden |22-26|. Dem Verfasser ist aber nur eine Untersuchung |24| bekannt, die sich mit der DM und der auch in dieser Arbeit beabsichtigten Adaption des Rekonstruktionsnetzwerkes an das Sprachsignal beschäftigt. In |24| wurde die Adaption mit Hilfe des Sprachsignals selber vorgenommen. Dabei gehen aber einige Vorteile der DM verloren. Man muß jetzt das Signal vor der eigentlichen DM A/D wandeln, um jeweils neue Koeffizienten für das Rekonstruktionsnetzwerk berechnen zu können. Diese Koeffizienten müssen dann zusätzlich zum Empfänger übertragen werden. Mit diesem Verfahren wurden gewisse Erfolge erzielt. Bei Rekonstruktionsnetzwerken höheren Grades aber wurde das

System instabil. Die Gründe dafür sollen in dieser Arbeit angegeben werden. Ein Versuch in |24|, die Adaption des Rekonstruktionsnetzwerkes mit einem Gradientenverfahren mit Hilfe des rekonstruierten Signals selber durchzuführen, führte zu nicht befriedigenden Ergebnissen.

In den Arbeiten, die sich mit der Adaption des Rekonstruktionsnetzwerkes bei der DPCM beschäftigen, wurde dagegen über recht befriedigende Ergebnisse bei der Anwendung von Gradientenverfahren berichtet.

Die Gradientenverfahren sind iterative Verfahren, bei denen Konvergenzprobleme auftreten können. Aus diesem Grund wird in dieser Arbeit versucht, ein direktes Verfahren zu finden, bei dem zusätzlich der Rechenaufwand nicht zu groß wird, so daß das Rekonstruktionsnetzwerk zu jedem Abtastzeitpunkt neu optimiert werden kann. Neben der Güte des rekonstruierten Signals ist die Gewährleistung der Stabilität des Deltamodulationsprozesses sehr wichtig. Ein Deltamodulationssystem gehört zur Klasse der nichtlinearen Abtastregelungssysteme |27|. In einem solchen System können, neben Stabilitätsproblemen im linearen Teil, Grenzzyklusschwingungen auftreten. Es werden Bedingungen dafür angegeben, daß abgesehen von solchen Grenzzyklusschwingungen, die nicht stören, mit großer Sicherheit keine anderen Grenzzyklusschwingungen auftreten können.

Da die Entwicklung und Optimierung des in dieser Arbeit untersuchten Deltamodulators wesentlich auf Ergebnissen aus der Rate Distortion Theorie und den Modellvorstellungen des Sprachprozesses beruht, werden zunächst diese beiden Gebiete betrachtet. Bei der Rate Distortion Theorie ist im Zusammenhang mit der DM die Auswirkung der Überabtastung der Signalquelle auf die zu erreichende Rekonstruktionsgüte von besonderer Bedeutung.

2. Ergebnisse der Rate Distortion Theorie (RDT)

Die Aufgabe eines Nachrichtenübertragungssystems (Bild 2.1) ist die effiziente und möglichst fehlerfreie Übertragung von Information von einer Signalquelle† zum Empfänger. Im Rahmen

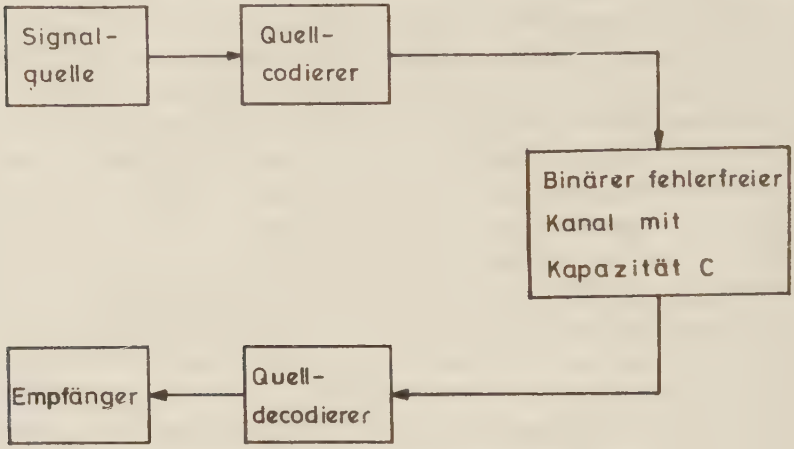


Bild. 2.1 Nachrichtenübertragungssystem

dieser Arbeit werden von diesem Übertragungssystem nur der Quellcodierer bzw. -decodierer betrachtet.

Nach Shannon |28| kann man dem Kanal eine gewisse Kapazität C (z.B. in bit/s) zuordnen. Die Realisierung dieses Kanals soll dabei keine Rolle spielen.

Damit die Information einer Signalquelle, mit der Entropie H, mit beliebig kleiner Fehlerhäufigkeit über den Kanal übertragen werden kann, muß gelten |19|:

$$H \cdot L \leq C \quad . \tag{2.1}$$

Dabei werden

H in bit/Symbol (Letter)

und L in Symbole/s

gemessen.

† Die statistischen und informationstheoretischen Kenngrößen von wertekontinuierlichen Zufallsprozessen werden im Anhang A1 definiert.

Bei einer wertediskreten Signalquelle läßt sich die Entropie H mit Hilfe der wertediskreten Wahrscheinlichkeitsverteilung der Signalwerte bestimmen, und man kann dann mit der Bedingung (2.1) nachprüfen, ob diese Information fehlerfrei übertragen werden kann.

Bei wertekontinuierlichen Signalquellen ist ein solches Vorgehen bei einem fehlerfreien Kanal nicht möglich, da die Entropie einer kontinuierlichen Signalquelle unendlich groß ist. In der Praxis bedeutet das, daß man bei der Übertragung eines Signalwertes einen gewissen Fehler zulassen muß.

Mit Hilfe der RDT läßt sich nun bei kontinuierlichen Signalquellen ebenfalls eine der Größe H•L entsprechende Informationsrate (rate) R berechnen, wenn ein bestimmtes mittleres Maß an Verzerrung (distortion) D zugelassen werden darf oder vom Empfänger der Information toleriert wird. Dabei schreibt der Empfänger das Verzerrungsmaß vor.

Aufgrund dieses Verzerrungsmaßes und der statistischen Beschreibung der Signalquelle ist es möglich, die Informationsrate der Signalquelle im Bezug auf das Verzerrungsmaß zu berechnen. Die so berechnete Rate einer kontinuierlichen Signalquelle bekommt ihre volle Bedeutung dadurch, daß man in der Lage ist, eine Codierung der Quelle so vorzunehmen (Quellcodierungstheorem|19|, daß die Rate des erzeugten Quellcodes beliebig genau mit der berechneten Rate der Quelle übereinstimmt.

Die RDT läßt sich aber ebensogut bei wertediskreten Signalquellen anwenden, da in der Praxis der Übertragungskanal nicht fehlerfrei ist. Bei geeigneter Beschreibung des Fehlerverhaltens des Kanals und der Vorgabe eines Fehlerbewertungsmaßes, läßt sich dann auf ähnliche Weise für eine gegebene wertediskrete Siganlquelle berechnen, welchen mittleren Fehler man mindestens bei der Übertragung über diesen Kanal hinnehmen muß.

2.1 Charakterisierung von zeitdiskreten, stationären, gaußschen Signalquellen mit Gedächtnis

Für Signalquellen mit frequenzbandbeschränktem Leistungsspektrum†

$$S(\omega) = 0 \text{ für } |\omega| = |2\pi f| > \omega_o = 2\pi f_o = \frac{2\pi}{T_o}, \quad f=\text{Frequenz} \qquad (2.2)$$

läßt sich die Zeitfunktion x(t) eindeutig und fehlerfrei aus ihren Abtastwerten x(nT) rekonstruieren, wenn die Abtastrate F=1/T größer als die zweifache Signalbandbreite f_o ist (Abtasttheorem). Für Zufallsfolgen, die eine solche Signalquelle erzeugt, sollen folgende Definitionen gelten:

$$x(nT) = x_n \qquad \text{Abtastwert,} \qquad (2.3)$$

$$\underline{x} = [x_n, x_{n-1}, \ldots, x_{n-N+1}]^T \qquad \text{Signalvektor (transponiert) der Länge N,} \qquad (2.4)$$

$$\overline{\underline{x}} = E\{\underline{x}\} = \underline{0} \quad \text{o.B.d.A.} \qquad \text{Erwartungswert von } \underline{x}, \qquad (2.5)$$

$$\underline{R}_x = E\{\underline{x} \cdot \underline{x}^T\} = \begin{bmatrix} r_x(0) & \cdots & r_x([N-1]T) \\ \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ r_x([N-1]T) & \cdots & r_x(0) \end{bmatrix} \qquad \text{Kovarianzmatrix} \qquad (2.6)$$

mit der Determinante $|R_x|$.

Für stationäre Signalquellen gilt, daß die Kovarianzmatrix $\underline{R}_x$ Töplitzform besitzt und alle Eigenwerte

$$\lambda_k^{(N)}, \quad k = 1,2,\ldots, N \qquad (2.7)$$

positiv sind. Wenn die Korrelationswerte außerhalb der Hauptdiagonalen von Null verschieden sind, handelt es sich um Quellen mit Gedächtnis. Wenn es sich um einen gaußschen Zufallsprozeß handelt, dann ist die Wahrscheinlichkeitsdichtefunktion (WDF) $g(\underline{x})$ dieses Prozesses vollkommen durch die Angabe der Kovarianzmatrix bestimmt:

$$g(\underline{x}) = (2\pi)^{-\frac{N}{2}} |\underline{R}_x|^{-\frac{1}{2}} \exp\left[-\tfrac{1}{2}\, \underline{x}^T \cdot \underline{R}_x^{-1} \cdot \underline{x}\right] . \qquad (2.8)$$

---

†Es wird diese Schreibweise gewählt, da S(ω) als Fouriertransformierte der reellen und geraden Zeitfunktion $r_x(t)$ (Autokorrelationsfunktion von x(t)) ebenfalls eine reelle und gerade Funktion in ω ist.

2.2 Definition der Rate Distortion Funktion (RDF) und Variationsproblem zur Berechnung der RDF für eine Signalquelle ohne Gedächtnis

Gegeben sind die Signalquelle {X,p(x)}, der "Empfänger" {Y,q(y)} und ein Verzerrungsmaß $\rho(x,y)$. Dann lassen sich folgende Beziehungen zwischen Quelle und Empfänger definieren:

Erwartungswert der Verzerrung

$$d = E\{\rho(x,y)\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p_v(x,y)\ \rho(x,y)\ dx\ dy \quad , \tag{2.9}$$

Transinformation†

$$T(x;y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p_v(x,y)\ \log\left[\frac{p_v(x,y)}{p(x)q(y)}\right] dx\ dy \quad , \tag{2.10}$$

mit $$p(x) = \int_{-\infty}^{+\infty} p_v(x,y)\ dy \quad , \tag{2.11}$$

$$q(y) = \int_{-\infty}^{+\infty} p_v(x,y)\ dx \quad . \tag{2.12}$$

Dabei ist $p_v(x,y)$ die Verbund-WDF zwischen x und y. Die Beziehungen (2.9), (2.10) und (2.11) lassen sich mit

$$\begin{aligned} p_v(x,y) &= p(x)\ q_b(y/x) \quad , \\ q_b(y/x) &= \text{bedingte WDF}^{\dagger\dagger} \quad , \end{aligned} \tag{2.13}$$

folgendermaßen umformulieren:

$$d\ [q_b(y/x)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x)\ q_b(y/x)\ \rho(x,y)\ dx\ dy \quad , \tag{2.14}$$

$$T\ [q_b(y/x)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x)\ q_b(y/x)\ \log\left[\frac{q_b(y/x)}{q(y)}\right] dx\ dy, \tag{2.15}$$

mit $$q(y) = \int_{-\infty}^{+\infty} p(x)\ q_b(y/x)\ dx \quad . \tag{2.16}$$

† Die Basis der Logarithmierung bestimmt die Maßeinheit der Information. Durch die Verwendung von $\log(\cdot)$ statt $\ln(\cdot)=\log_e(\cdot)$ oder $\mathrm{ld}(\cdot)=\log_2(\cdot)$ soll diese Maßeinheit noch offen bleiben.

†† Siehe Anhang A1.

Die RDF R(D) ist nun definiert als:

$$R(D) = \inf_{q_b \in Q_D} \{T[q_b(y/x)]\} \tag{2.17}$$

mit $Q_D := \{q_b(y/x): d[q_b] = D\}$. (2.18)

R(D) ist also die minimale Transinformation zwischen {X} und {Y} unter der Nebenbedingung, daß der mittlere Rekonstruktionsfehler $d[q_b]$ den vorgeschriebenen Wert D besitzt. Die bedingte WDF $q_b(y/x)$ kennzeichnet dabei die statistischen Übertragungseigenschaften des Systems. Dabei müssen zwei weitere Nebenbedingungen von $q_b(y/x)$ eingehalten werden:

$$\int_{-\infty}^{+\infty} q_b(y/x)dy = 1 \qquad \forall\ x \tag{2.19}$$

und $q_b(y/x) \geq 0$. (2.20)

Die einzige Annahme, die gemacht werden muß, um zu überprüfen, daß R(D) die effektive Informationsrate der Signalquelle {X,p(x)} ist, ist die Existenz eines Wertes $\tilde{y}$, so daß gilt:

$$\int_{-\infty}^{+\infty} p(x)\rho(x,\tilde{y})dx < \infty \quad . \tag{2.21}$$

Zur Lösung dieses Variationsproblems kann es zweckmäßig sein, [19] zunächst die Nebenbedingung (2.20) nicht zu beachten. Nach der Lösung dieses so abgeänderten Variationsproblems ist diese Lösung nur in dem Bereich von D-Werten gleich R(D), in dem die Bedingung (2.20) eingehalten wird.

Der interessierende D-Bereich ist bei kontinuierlichen Signalquellen durch

$$0 < D < D_{max} \tag{2.22}$$

gegeben. Die Verzerrung D = 0 läßt sich nicht erreichen, da hierzu eine unendliche Informationsrate notwendig ist. Der Wert D_{max} ist dadurch definiert, daß von diesem Wert ab die Informationsrate der Signalquelle gleich Null wird:

$$R(D) = 0 \qquad \text{für} \qquad D \geq D_{max}$$

mit $D_{max} = \inf_{y} \int_{-\infty}^{+\infty} p(x)\rho(x,y)dx$.

Im übrigen ist R(D) im Bereich (2.22) eine nichtnegative monoton fallende, konvexe Funktion von D |19|. Dieser grundsätzliche Verlauf von R(D) ist in Bild 2.2 dargestellt.

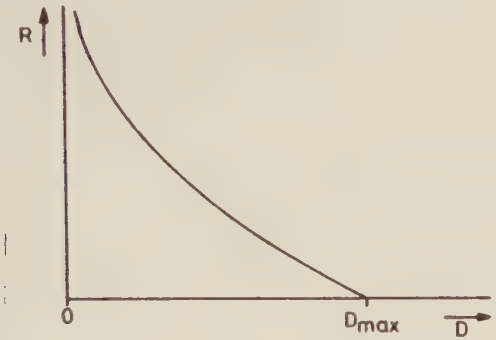


Bild 2.2 Grundsätzlicher Verlauf von R(D)

2.3 R(D) für stationäre, gedächtnisfreie, gaußsche Signalquellen mit dem Verzerrungsmaß mittlerer quadratischer Fehler (MQF)

Eine Schwierigkeit bei der Bestimmung der RDF besteht in der Auswahl eines geeigneten Verzerrungsmaßes $\rho(x,y)$ durch den Empfänger. Dieses Maß muß gleichzeitig physikalisch (physiologisch) sinnvoll und mathematisch beschreibbar sein. Bei der Beurteilung der Rekonstruktion z.B. eines Sprachsignales ist man letztlich auf eine subjektive vom jeweiligen Hörer abhängige Bewertung angewiesen. Solche Gütemaße lassen sich aber kaum mathematisch darstellen.

Ein Maß, welches mathematisch beschreibbar und physikalisch sinnvoll ist, ist der quadratische Fehler zwischen x und y, d.h.:

$$\rho(x,y) = (x-y)^2 \quad . \qquad (2.23)$$

In diesem Fall ist D der MQF zwischen x und y bzw. die Geräuschleistung des Fehlersignals. Bei ergodischen Prozessen kann man diese Geräuschleistung auch als den zeitlichen Mittelwert des quadratischen Fehlers bestimmen:

$$D = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{+T/2} [x(t) - y(t)]^2 \, dt \quad . \tag{2.24}$$

In einem Nachrichtenübertragungssystem ist diese Größe also leicht zu messen. Für ein solches Verzerrungsmaß und für eine gaußsche Signalquelle ohne Gedächtnis (p(x)=g(x)) mit

$$g(x) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[-\frac{x^2}{2\sigma^2}\right]$$

und $\sigma^2 = r_x(0)$, $\overline{x} = 0$,

erhält man als RDF |19|:

$$R(D) = \begin{cases} \frac{1}{2} \log \frac{\sigma^2}{D} & \text{für } 0 \le D \le \sigma^2 \\ 0 & \text{für } D \ge \sigma^2 \end{cases} \quad . \tag{2.25}$$

Dabei ergeben sich:

$$q(y) = \left[2\pi(\sigma^2-D)\right]^{-\frac{1}{2}} \exp\left[-\frac{y^2}{2(\sigma^2-D)}\right] \quad , \tag{2.26}$$

$$q_b(y/x) = (2\pi\beta D)^{-\frac{1}{2}} \exp\left[-\frac{(y-\beta x)^2}{2\beta D}\right] \tag{2.27}$$

mit $\beta = 1 - \frac{D}{\sigma^2}$. (2.28)

Weiterhin gilt folgender wichtiger Satz (Theorem 4.3.2 in |19|):

Satz 2.1 Sei {X} eine Quelle mit der WDF p(x), $\overline{x}$=0 und Leistung σ^2, dann gilt für diese Quelle mit $\rho(x,y) = (x-y)^2$

$$R(D) \le \frac{1}{2} \log \frac{\sigma^2}{D} \quad . \tag{2.29}$$

Die RDF für eine gaußsche Signalquelle ohne Gedächtnis ist also eine obere Schranke für alle Signalquellen mit gleicher Leistung, wenn D der MQF ist.

In dieser Arbeit wird auch häufig ein logarithmisches Maß für den MQF benutzt, und zwar der Störabstand SNR (Signal-to-noise ratio) bzw. das logarithmische Maß für den Kehrwert NSR:

$$SNR = 10 \log_{10}\left[\frac{\sigma^2}{D}\right] \text{ in dB} \quad , \tag{2.30a}$$

$$NSR = 10 \log_{10}\left[\frac{D}{\sigma^2}\right] \text{ in dB} \quad . \tag{2.30b}$$

2.4 R(D) für gaußsche Signalquellen mit Gedächtnis

Wenn man eine Signalquelle ohne Gedächtnis codiert, so ist jeder Signalwert nur entsprechend seiner Wahrscheinlichkeit zu betrachten. Bei Signalquellen mit Gedächtnis hat man aber schon eine gewisse Information über jeden neuen Signalwert durch die vorangehenden Signalwerte. Es ist also wichtig zu wissen, um welchen Betrag sich die Informationsrate R(D) der Quelle bei konstantem D gegenüber einer Signalquelle ohne Gedächtnis vermindert. Dadurch läßt sich bei solchen Signalquellen eine entsprechende Datenkompression erreichen.

Die Berechnung von R(D) für Quellen mit Gedächtnis ist sehr schwierig. Bis heute sind exakte Lösungen nur für gaußsche Signalquellen mit dem Gütekriterium MQF und für einige spezielle nichtgaußsche Signalquellen bekannt. Wie bei Signalquellen ohne Gedächtnis kommt den gaußschen Signalquellen eine gewisse Sonderstellung zu, da mit ihrer Hilfe eine Schranke für die zu erreichende Güte für alle anderen Signalquellen angebbar ist.

Die RDF für Signalquellen mit Gedächtnis $\{X, p(\underline{x})\}$ ist folgendermaßen definiert:

$$R(D) = \lim_{N\to\infty} R_N(D) \tag{2.31}$$

$$\text{mit} \quad R_N(D) = \frac{1}{N} \inf_{q_b \in Q_D} \{T[q_b(\underline{y}/\underline{x})]\} \quad , \tag{2.32}$$

$$Q_D := \{q(\underline{y}/\underline{x}): d(q_b) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(\underline{x}) q_b(\underline{y}/\underline{x}) \rho_N(\underline{x},\underline{y}) \underline{dx}\,\underline{dy} < D\} \quad , \tag{2.33}$$

$$\rho_N(\underline{x},\underline{y}) = \frac{1}{N} \sum_{k=1}^{N} \rho(x_k,y_k) \quad , \tag{2.34}$$

$$T[q_b(\underline{y}/\underline{x})] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p(\underline{x}) q_b(\underline{y}/\underline{x}) \log\left[\frac{q_b(\underline{y}/\underline{x})}{q(\underline{y})}\right] \underline{dx}\ \underline{dy} \quad , \tag{2.35}$$

$$q(\underline{y}) = \int_{-\infty}^{+\infty} p(\underline{x})\ q_b(\underline{y}/\underline{x})\ \underline{dx} \quad . \tag{2.36}$$

$$q_b(\underline{y}/\underline{x}) \geq 0 \quad . \tag{2.37}$$

Jedes Integral ist dabei N-fach zu betrachten. Bei gaußschen Signa
quellen gelangt man mit Hilfe einer unitären Transformation des
Signalvektors $\underline{x}$ zu N statistisch unabhängigen Teilquellen x'_k,
$k = 1,2,\ldots,N$:

$$\underline{x}' = \underline{\Gamma}^{-1}\underline{x} \tag{2.38a}$$

bzw.

$$y' = \underline{\Gamma}^{-1}\underline{y} \quad . \tag{2.38b}$$

Dabei ist die Transformationsmatrix $\underline{\Gamma}$ definiert durch:

$$\underline{R}_x = \underline{\Gamma} \begin{bmatrix} \lambda_1^{(N)} & & 0 \\ & \ddots & \\ 0 & & \lambda_N^{(N)} \end{bmatrix} \underline{\Gamma}^{-1} \quad . \tag{2.39}$$

Die $\lambda_k^{(N)}$ sind dabei die in (2.7) definierten Eigenwerte von $\underline{R}_x$.
Durch diese Transformation erhält man eine Produktdarstellung
der gauß'schen WDF für N statistisch unabhängige Teilquellen:

$$g(\underline{x}') = \prod_{k=1}^{N} (2\pi\lambda_k^{(N)})^{-\frac{1}{2}} \exp\left[-\frac{x_k'^2}{2\lambda_k^{(N)}}\right] \quad . \tag{2.40}$$

Da die Transinformation unabhängig von einer Maßstabsänderung
ist, folgt daraus, daß die RDF jetzt als arithmetischer Mittel-
wert der Rate Distortion Funktionen der Teilquellen ermittelt
werden kann. Zusätzlich bleibt durch die unitäre Transformation
der Kovarianzmatrix auch die Euklidsche Fehlernorm erhalten.
Da es sich jetzt um N statistisch unabhängige Signalquellen
handelt, erhält man auch den MQF D als arithmetischen Mittel-

wert der mittleren quadratischen Fehler, die bei der Rekonstruktion der Teilquellen entstehen. Bei der Vorgabe eines MQF's θ, der bei der Rekonstruktion jeder einzelnen Teilquelle zugelassen werden darf, erhält man schließlich die RDF für eine gaußsche Signalquelle in Abhängigkeit vom Parameter θ :

$$R_N(D) = \frac{1}{N} \sum_{k=1}^{N} \max(0, \frac{1}{2} \log \frac{\lambda_k^{(N)}}{\theta}) \quad , \tag{2.41}$$

$$D(\theta) = \frac{1}{N} \sum_{k=1}^{N} \min(\theta, \lambda_k^{(N)}) \quad . \tag{2.42}$$

Dieses Ergebnis läßt sich mit Hilfe von (2.25) einsehen, wenn man beachtet, daß bei einem MQF θ die Teilquelle k keinen Beitrag mehr zur Information der Gesamtquelle liefert, wenn die Leistung $\lambda_k^{(N)}$ der Teilquelle k kleiner als θ ist.

Mit (2.41) und (2.42) besteht die Möglichkeit $R_N(D)$ numerisch zu ermitteln. Man berechnet zunächst Schätzwerte für die Autokorrelationswerte des betrachteten gaußschen Prozesses. Zur Berechnung der Eigenwerte von symmetrischen Matrizen existieren leistungsfähige Rechenalgorithmen [29]. Durch Betrachtung von Kovarianzmatrizen mit wachsender Ordnung N läßt sich dann der Grenzwert für R(D) (2.31) abschätzen.

Da das Leistungsspektrum des betrachteten gaußschen Prozesses nach Abschnitt 2.1 (2.2) frequenzbandbegrenzt sein soll, läßt sich auch noch eine andere Darstellung von (2.41) und (2.42) für den Grenzwert $N \to \infty$ finden. Die z-Transformierte S_{zx} des Leistungsspektrums erhält man aus den Abtastwerten der Autokorrelationsfunktion:

$$S_{zx}(z) = \sum_{n=-\infty}^{+\infty} r_{xk} z^{-k} \tag{2.43}$$

mit $\quad r_{xk} = r_x(kT) \quad .$

Mit Hilfe eines Satzes über unendliche Töplitzmatrizen [19,30] ist man in der Lage, in (2.41) und (2.42) den Grenzübergang für $N \to \infty$ vorzunehmen. Man erhält mit $z = e^{j\varphi}$:

$$R(D) = \frac{1}{4\pi} \int_{-\pi}^{+\pi} \max\left[0, \log \frac{S_{zx}(\varphi)}{\theta}\right] d\varphi \quad , \tag{2.44}$$

$$D(\theta) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \min\left[\theta, S_{zx}(\varphi)\right] d\varphi \quad . \tag{2.45}$$

Eine anschauliche Darstellung zur Bestimmung von $D(\theta)$ gibt Bild 2.3:

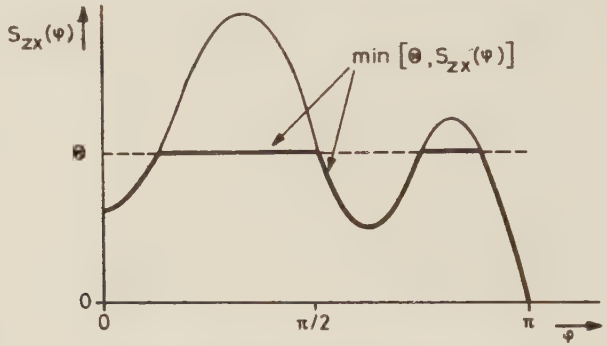


Bild 2.3 Bestimmung von $D(\theta)$ für $0 \leq \varphi \leq \pi$

Die dick ausgezogene Kurve

$$G(\varphi) = \min\left[\theta, S_{zx}(\varphi)\right] \tag{2.46}$$

stellt dabei das Leistungsspektrum des Rekonstruktionsfehlers dar, das im optimalen Fall erreicht wird.

Die oben angegebene parametrische Darstellung der RDF für stationäre gaußsche Signalquellen mit Gedächtnis besitzt eine weitere wichtige Eigenschaft entsprechend Satz 2.1. Es gilt nämlich (Theorem 4.6.3 in [19]):

Sei $\{X\}$ eine zeitdiskrete Signalquelle mit stationärer Kovarianzmatrix, Mittelwert Null und Leistungsspektrum $S_{zx}(\varphi)$ d.h.:

$$E\{x_n\} = 0 \quad ,$$

$$E\{x_n x_{n+k}\} = r_{xk} = \frac{1}{2\pi} \int_{-\pi}^{+\pi} S_{zx}(\varphi)\, e^{jk\varphi} d\varphi \quad ,$$

dann gilt für die RDF dieser Signalquelle:

$$R(D) \leq R_G(D) \quad . \tag{2.47}$$

Dabei ist $R_G(D)$ durch die parametrische Darstellung (2.44) und (2.45) festgelegt. Das bedeutet, daß von allen Signalquellen mit gleichem Leistungsspektrum eine gaußsche Signalquelle die größte Informationsrate R_G bei gleichem zulässigem MQF D besitzt.

2.5 R(D) für überabgetastete Signalquellen bei vorgegebener Kanalkapazität C

Will man bei der Codierung einer Signalquelle mit dem Leistungsspektrum $S_{zx}(\varphi)$ bei fester vorgegebener Kanalkapazität C und einer Signalbandbreite f_o die Kanalkapazität voll ausnutzen, so kann man die Signalquelle mit $B=C/2f_o$ bit/Abtastwert bei Abtastung mit der Nyquistrate $2f_o$ codieren. Bei der Wahl einer höheren Abtatsrate $F=M2f_o$ ($M \geq 1$) wird die Zahl der pro Abtastwert zur Verfügung stehenden bits B um den Faktor M kleiner. Gleichzeitig wird aber auch die Korrelation zwischen den Abtastwerten größer, so daß auch die Information pro Abtastwert für einen bestimmten MQF D kleiner wird.

Wie sich die soeben beschriebenen Verhältnisse ändern, soll im folgenden an einem numerischen Beispiel verdeutlicht werden. Es wird eine Signalquelle mit dem logarithmischen Leistungsspektrum

$$LS(f) = 10 \log_{10} [S_{zx}(f)] \text{ in dB} \tag{2.48}$$

mit

$$\varphi = \frac{2\pi f}{F}$$

nach Bild 2.4 betrachtet. Das ist das Leistungsspektrum eines Sprachlautes, der in Abschnitt 6 noch weiter untersucht wird.

Bei der Berechnung der RDF'en für verschiedene Abtastraten (Bild 2.5) wird eine normierte Darstellung mit dem Normierungsfaktor F/C verwendet. Der Wert R·F/C gibt dann den Ausnutzungsfaktor des Kanals an, der im Bereich

$$0 \leq RF/C \leq 1 \tag{2.49}$$

liegen muß. Bei einer in diesem Beispiel zugrunde gelegten Kanal-

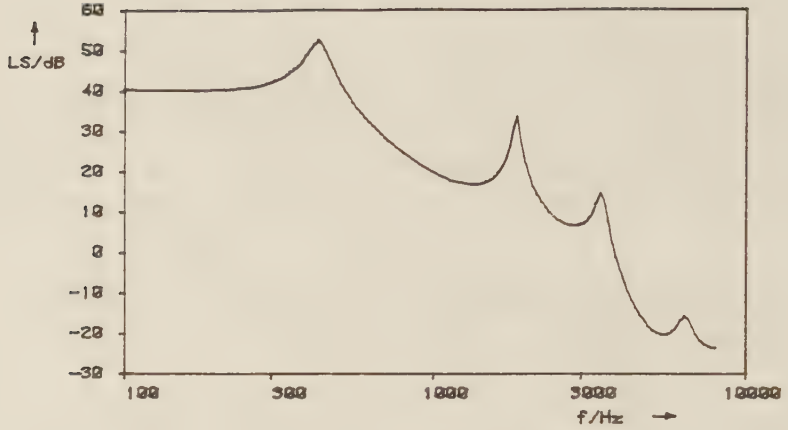


Bild 2.4 Log. Leistungsspektrum LS(f)

kapazität von 32 kbit/s sind die zu erreichenden Rekonstruktionsgüten bei voller Kapazitätsausnutzung recht stark von der Abtastrate abhängig (für das gegebene Leistungsspektrum der Signalquelle). Bei niedrigeren Kanalkapazitäten dagegen ist der Unterschied nicht mehr so groß, wie an einem Beispiel in Abschnitt 4.3 gezeigt werden wird. Aus diesem Grunde sollte man die Abtastrate so niedrig wie möglich wählen. Dabei stellt sich natürlich die Frage, ob dann bei der Deltamodulation mit einer 1-bit Quantisierung pro Abtastwert die zu erreichende Rekonstruktionsgüte noch ausreichend ist.

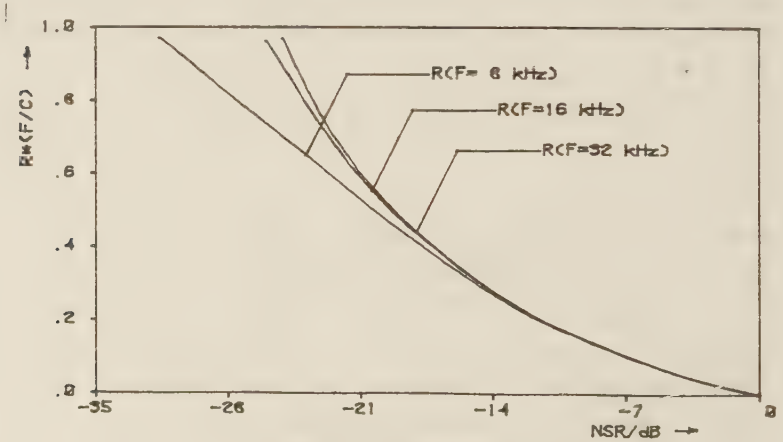


Bild 2.5 R(F,NSR)*(F/C)/Kanalkapazitaet C=32 kbit/s
Gedaechtnislaenge= .6 ms

2.6 Autoregressive Signalquellen

Im Zusammenhang mit der Codierung von Sprachsignalen haben sogenannte autoregressive Signalquellen eine große Bedeutung. Solche Signalquellen stellen gute Modellquellen für den Sprachprozeß dar |20,21|. Darüber hinaus läßt sich für gaußsche autoregressive Signalquellen die RDF berechnen und es existiert ein Quellcodierungstheorem |19|, das garantiert, daß sich solche Quellen mit einem Code darstellen lassen, dessen Informationsrate beliebig nahe bei R(D) liegt. Im folgenden werden solche Signalquellen vorgestellt und es wird diskutiert, wie dieser Code erzeugt werden kann. Daraus ergeben sich wichtige Schlußfolgerungen für die Erzeugung eines möglichst optimalen Codes durch einen Deltamodulator.

2.6.1 Definition und RDF von gaußschen , autoregressiven Signalquellen

Eine autoregressive Signalquelle {X} wird durch folgende Differenzengleichung beschrieben:

$$x_n = e_n - \sum_{k=1}^{M} a_k x_{n-k} \quad . \tag{2.50}$$

$\{e_n\}$ ist dabei eine zeitdiskrete Folge von unabhängigen Zufallsvariablen mit gleicher WDF p(e). Im Falle einer gaußschen, autoregressiven Quelle ist $\tilde{p}(e) = g(e)$ nach Gleichung (2.8) mit N=1, $r_x(0)=\sigma_e^2$ und Mittelwert Null. Es läßt sich zeigen, daß $\{x_n\}$ dann ebenfalls eine Folge von gaußschen Zufallsvariablen ist, wobei die x_n dann im allgemeinen aber nicht mehr unkorreliert sind. Wenn alle Nullstellen des Polynoms

$$A(z) = \sum_{k=0}^{M} a_k z^{-k} \qquad \text{mit } a_o=1 \tag{2.51}$$

und der komplexen Variablen z innerhalb des Einheitskreises liegen, ergibt sich für das Leistungsspektrum von $\{x_n\}$:

$$S_{zx}(\varphi) = \frac{\sigma_e^2}{\left| \sum_{k=0}^{M} a_k e^{-jk\varphi} \right|^2} \quad . \tag{2.52}$$

Wie in [19] gezeigt wird, spielt der Anfangszustand

$$\underline{x}_o = [x_{-1}, x_{-2}, \ldots, x_{-M}]^T \tag{2.53}$$

bei der Bestimmung der RDF keine Rolle. Damit ist R(D) für (2.50), wenn D der MQF ist, durch (2.44) und (2.45) gegeben. Für $S_{zx}(\varphi)$ ist der Ausdruck (2.52) einzusetzen.

Wenn man $\underline{x}_o = \underline{0}$ setzt, kann man (2.50) auch in einer Matrixform darstellen:

$$\underline{e} = \underline{A}\,\underline{x}$$

mit

$$\underline{e} = [e_n, e_{n-1}, \ldots e_{n-N+1}]^T \tag{2.54}$$

und

$$\underline{A} = \begin{bmatrix} a_o & & & & \\ a_1 & & & 0 & \\ \cdot & a_1 & \cdot & & \\ \cdot & \cdot & \cdot & & \\ a_M & & \cdot & & \\ & \cdots & & a_1 & \\ 0 & & a_M \cdots & a_1 & a_o \end{bmatrix} . \tag{2.55}$$

2.6.2 Baumcodierung von gaußschen , autoregressiven Signalquellen

Es läßt sich zeigen, daß Baumcodes (Bild 2.6) existieren, mit denen man gaußsche, autoregressive Signalquellen beliebig gut im Sinne der RDT codieren kann |19|. Baumcodes

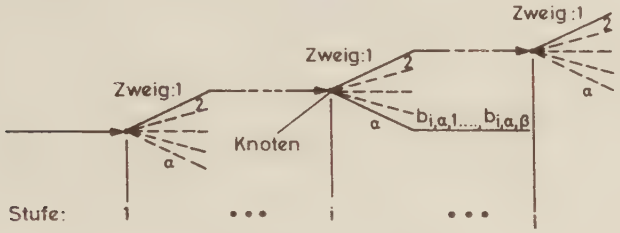


Bild 2.6 Baumcode mit: l Stufen
α Zweigen pro Knoten
β Symbolen b pro Zweig

haben die wichtige Eigenschaft, daß man bei der Erzeugung eines besten Codewortes für Blöcke von Signalwerten nicht unbedingt alle Codewörter des Codes überprüfen muß, sondern nur einen "Pfad" durch den Codebaum suchen muß, der das richtige Codewort angibt. Bei der Auswahl eines Pfades durch den Codebaum kann man verschiedene Strategien angeben, die aber alle zum Ziel haben, möglichst wenig Codewörter überprüfen zu müssen. Ein solcher Baumcode mit dem Codewortrepertoire

$$B = \{\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_\gamma\} \tag{2.56}$$

besitzt $K = \alpha^l$ (2.57)

Codewörter. Jedes Codewort hat die Blocklänge N=lβ. Die maximal verfügbare Informationsrate R_{CB} eines solchen Baumcodes beträgt

$$R_{CB} = \frac{1}{N} \operatorname{ld} K = \frac{1}{\beta} \operatorname{ld}\alpha \quad \text{bit/Symbol} \tag{2.58}$$

unabhängig von der Stufenzahl l. Für einen solchen Baumcode läßt sich folgender Existenzsatz zur Quellcodierung von gaußschen, autoregressiven Quellen angeben (Theorem 6.3.4 in |19|):

Satz 2.2 Es sei {X} eine nach (2.50) definierte gaußsche, autoregressive Signalquelle und R(D) die RDF, wenn D der MQF zwischen {X} und der Rekonstruktion {Y} ist, und es sei δ ein Wert mit der Eigenschaft

$$\delta = \inf_{\varphi} \frac{\sigma_e^2}{\left| \sum_{k=0}^{M} a_k e^{-jk\varphi} \right|^2} \quad . \tag{2.59}$$

Bei vorgegebenen Werten $\varepsilon > 0$ und $D \leq \delta$ existiert dann für hinreichend große Blocklängen N ein Baumcode für {X}, dessen maximal verfügbare Informationsrate kleiner ist als $R(D)+\varepsilon$ und für den der MQF ebenfalls kleiner ist als $D+\varepsilon$.

Zur Herleitung dieses Satzes muß man zeigen, daß die rekonstruierten Signalwerte $\underline{y} = [Y_n, y_{n-1}, \ldots, y_{n-N+1}]^T$ so gewählt werden können, daß für ihre WDF $q(\underline{y})$ Gleichung (2.36) gilt, wobei $q_b(\underline{y}/\underline{x})$ diejenige WDF ist, die das Variationsproblem zur Berechnung von R(D) löst. Weiterhin müssen die y_n die erforderliche Baumstruktur besitzen. Dieser zweite Teil der Herleitung soll nicht weiter betrachtet werden.

In diesem Zusammenhang ist es im wesentlichen nur interessant, den ersten Teil der Herleitung zumindest teilweise zu verfolgen. Aus der Kenntnis der Berechnung von $R_N(D)$ für gaußsche Signalquellen mit Gedächtnis in Abschnitt 2.4 ist bekannt, daß man die WDF $q(\underline{y})$ dadurch erhält, daß man eine unitäre Transformation von $\underline{x}$ und $\underline{y}$ mit der Transformationsmatrix $\underline{\Gamma}$ gemäß (2.38) durchführt, d.h.:

$$\underline{x}' = \underline{\Gamma}^{-1}\underline{x} \quad \text{und} \quad \underline{y}' = \underline{\Gamma}^{-1}\underline{y} \quad ,$$

um zu einer optimalen Codierung zu gelangen. Die optimale WDF der $\underline{y}'$ muß entsprechend (2.26) für eine gaußsche Quelle ohne Gedächtnis gebildet werden. Weiterhin kann man für gaußsche Signalquellen ohne Gedächtnis zeigen, daß man die optimale Codierung einer solchen Signalquelle durch einen sogenannten Rückwärtskanal beschreiben kann, wobei die Fehlerabtastwerte von den rekonstruierten Werten statistisch unabhängig sind. Dieser Rückwärtskanal hat für die transformierten Größen $\underline{x}'$, $\underline{y}'$ und $\underline{e}'$ folgende Form:

$$\underline{x}' = \underline{y}' + \underline{e}' \quad . \qquad (2.60)$$

Durch Rücktransformation folgt daraus:

$$\underline{x} = \underline{y} + \underline{\Gamma}\,\underline{e}' \quad . \qquad (2.61)$$

Bei autoregressiven Signalquellen ergibt sich weiter nach Multiplikation mit der Matrix $\underline{A}$ (2.55):

$$\underline{A}\,\underline{x} = \underline{e} = \underline{A}\,\underline{y} + \underline{A}\,\underline{\Gamma}\,\underline{e}' \quad . \qquad (2.62)$$

Dabei sind die Komponenten von $\underline{e}$ statistisch unabhängige gaußsche Zufallsvariable. Wenn man weiterhin definiert:

$$\underline{w} = \underline{e} - \underline{A}\,\underline{\Gamma}\,\underline{e}' \quad , \qquad (2.63)$$

dann ergibt sich für (2.62):

$$\underline{w} = [w_n, w_{n-1}, \cdots, w_{n-N+1}]^T = \underline{A}\,\underline{y} \quad . \qquad (2.64)$$

Die Differenzengleichung für optimale Werte y_n lautet also:

$$y_n = w_n - \sum_{k=1}^{M} a_k\, y_{n-k} \quad . \qquad (2.65)$$

Man erhält also optimale rekonstruierte Werte y_n, wenn man die Zufallsvariablen w_n durch das gleiche Netzwerk schickt, welches zur Erzeugung der Signalwerte x_n benutzt wird. Im Falle einer autoregressiven Quelle mit Gedächtnis sind aber die Zufallsvariablen w_n nicht statistisch unabhängig. Sie sind vielmehr durch Gleichung (2.63) bestimmt.

Um nun weiter zu zeigen wie die w_n erzeugt werden, bzw. wie das optimale Rekonstruktionsnetzwerk für die y_n insgesamt aufgebaut werden muß ist noch einmal darauf hinzuweisen, daß der Satz 2.2, der hier ja zum Teil erläutert werden soll, nur für $\theta \leq \delta$ gilt. Für diesen Fall ist aus Abschnitt 2.4 und mit (2.61) bekannt, daß sich der Prozeß {X} aus dem optimalen Prozeß {Y}und unabhängigem weißen Rauschen mit der Leistung $D = \theta$ zusammensetzt. Für das optimale Leistungsspektrum des rekonstruierten Signals gilt also für diesen Fall:

$$S_{zy,Opt}(z) = S_{zx}(z) - D \tag{2.66}$$

mit $$S_{zy,Opt}(z) = \sum_{k=-\infty}^{\infty} r_{yk} z^{-k} \quad , \; r_{yk} = r_y(kT).$$

Dabei ist $r_y(kT)$ die Autokorrelationsfunktion des im optimalen Fall rekonstruierten Signals y_n. Für eine autoregressive Signalquelle ist weiter $S_{zx}(z)$ durch (2.52) gegeben, so daß man schließlich erhält ($z = e^{j\varphi}$):

$$S_{zy,Opt}(\varphi) = \frac{\sigma_e^2 - D \left| \sum_{k=0}^{M} a_k \, e^{-jk\varphi} \right|^2}{\left| \sum_{k=0}^{M} a_k \, e^{-jk\varphi} \right|^2} \quad . \tag{2.67}$$

Man erhält also das optimale Leistungsspektrum, wenn man einen unkorrelierten gaußschen Prozeß $\{v_n\}$ mit der Leistung 1 und Mittelwert Null durch ein Netzwerk

$$H(z) = \frac{B(z)}{A(z)}$$

mit $$|B(z)|^2 = \sigma_e^2 - D|A(z)|^2 \tag{2.68}$$

und $$B(z) = \sum_{k=0}^{M} b_k \, z^{-k} \tag{2.69}$$

schickt. Dieses Netzwerk ist in Bild 2.7 dargestellt.

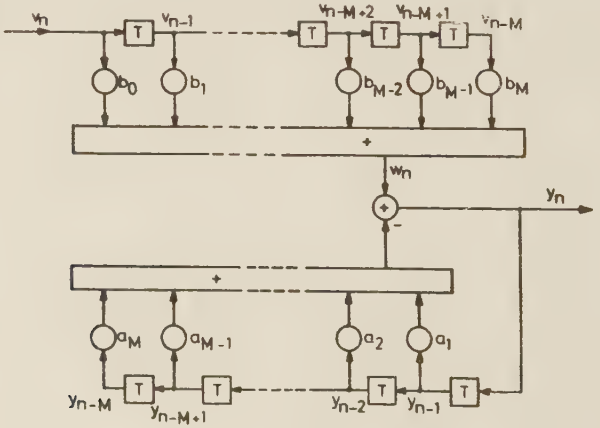


Bild 2.7 Optimales Rekonstruktionsnetzwerk für eine gaußsche, autoregressive Signalquelle

Da der Ausdruck auf der rechten Seite von (2.68) für $\Theta < \delta$ positiv ist, läßt sich zeigen, daß sich dieser Ausdruck durch das Quadrat des Betrages eines Polynoms B(z) nach (2.69) darstellen läßt.

Für sehr kleine Werte von D ist das Rekonstruktionsnetzwerk dabei nahezu identisch mit dem Filter, das den autoregressiven Prozeß erzeugt. Eine Baumcodierung erfolgt jetzt so, daß man die Zweige eines Codebaums, mit $R_{CB} = \frac{1}{\beta} \log \alpha = R(D)$, für $D \leq \delta$ mit gaußschen Zufallsvariablen aus einer Verteilung mit der Varianz 1 und dem Mittelwert Null besetzt. Für eine gegebene Blocklänge N sucht man jetzt denjenigen "Pfad" durch den Codebaum, dessen Codewörter die Abtastwerte für diesen Block am besten approximieren. Zwischen aufeinander folgenden Blöcken werden die Speicher des Netzwerks auf ihre Anfangswerte zurückgesetzt, damit die Codierung der Blöcke unabhängig voneinander erfolgt und so die maximal verfügbare Informationsrate des Codebaums erreicht wird.

Satz 2.2 gewährleistet dann, daß man die Informationsrate

$$R_{CB} = \frac{1}{N} \log K < R(D) + \varepsilon$$

so wählen kann, daß bei hinreichend großer Blocklänge N ein Baumcode existiert, bei dem der MQF < D+ε wird, bei beliebigem ε>0. Wenn man aber nur eine geringe Kanalkapazität zur Verfügung hat, so daß $\delta < \Theta$ gilt, dann gilt Satz 2.6 zunächst nicht mehr. Das Leistungsspektrum $S_{zy,Opt}$ zerfällt dann nach Maßgabe von $\Theta > \delta$ in mehrere Bänder (Bild 2.8) d.h. es muß nur noch

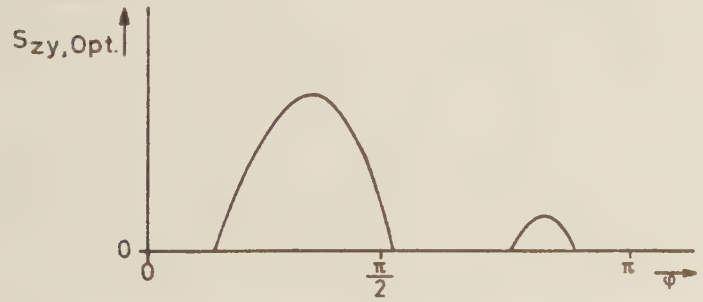


Bild 2.8 Optimales Leistungsspektrum $S_{zy,Opt.}$ für $\Theta > \delta$

das in diesen Frequenzbereichen enthaltene Signal übertragen werden.

Ein Netzwerk, welches einen solchen Frequenzgang besitzt, läßt sich mit einer endlichen Zahl von Verzögerungsgliedern nicht mehr realisieren. Eine beliebig gute Approximation ist aber immer noch möglich.

Ein wichtiges Ergebnis bei dieser Betrachtung der optimalen Baumcodierung ist, daß eine solche Codierung nur mit einer großen Blocklänge, d.h. mit einer großen Verzögerung möglich ist. Das gilt auch für die hier nicht betrachteten allgemeinen Blockcodes. Das Problem der Realisierung eines Baumcodierungsalgorithmus wurde ebenfalls nicht untersucht. Es muß zumindest die Verzerrung berechnet werden , die entsteht, wenn alle zur Konkurrenz zugelassenen Codewörter betrachtet werden. Dazu müssen dem Algorithmus die Abtastwerte der Signalquelle beliebig genau bekannt sein (im Rahmen der zulässigen Verzerrung).

Die Realisierung einer optimalen Baumcodierung erscheint also schon für stationäre Quellen sehr aufwendig zu sein. Für nichtstationäre Signalquellen (mit sich langsam ändernden statistischen Eigenschaften) müßte jeweils das optimale Rekonstruktionsfilter adaptiv berechnet werden. Auf dieses Problem der Realisierung einer optimalen Baumcodierung wird noch später, im Vergleich mit der in dieser Arbeit betrachteten DM mit adaptivem Rekonstruktionsnetzwerk, eingegangen.

3. Beschreibung des Sprachprozesses

Digitale Modulationsverfahren wie die DM gehören zur Klasse der Quellcodierungsverfahren, die den zeitlichen Verlauf z.B. vom Sprachsignal codieren. In Bild 3.1 ist der Zeitverlauf x(t) der akustischen Wellenform, wie sie durch die Ausgangsspannung eines Mikrophons repräsentiert wird, dargestellt†. Aus diesem Bild ist zumindest zu erkennen, daß der

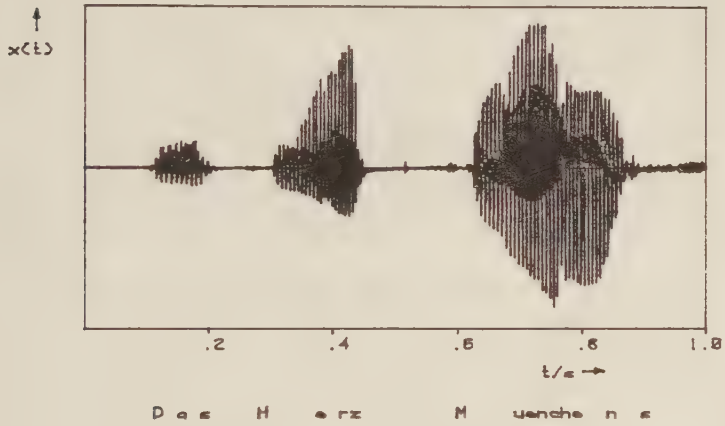


Bild 3.1 Zeitsignalverlauf von Sprache(M116M)

Sprachprozeß in unterschiedlichen Zeitsegmenten eine unterschiedliche, annähernd konstante Einhüllende besitzt. Betrachtet man solche Zeitabschnitte getrennt, so erkennt man, daß der Sprachprozeß dort unterschiedliche aber jeweils annähernd konstante statistische Eigenschaften aufweist. Man kann für solche Segmente die Autokorrelationsfunktion $r_x(nT) = r_{xn}$ (siehe auch Gleichung (3.10) in Abschnitt 3.2) für kurze Integrationszeiten (10 ....20 ms) bestimmen. Bild 3.2 zeigt das Ergebnis einer solchen Berechnung für einen Ausschnitt aus Bild 3.1.

Wenn man das Leistungsspektrum $S_{zx}(z)$ zu den jeweiligen Autokorrelationsfunktionen betrachtet, so kann man dieses Spektrum durch folgende Modellgleichung beschreiben:

$$S_{zx}(z) = |E(z)|^2 \cdot |H(z)|^2 \qquad (3.1)$$

† Zur Beschreibung der verwendeten Sprachproben siehe Anhang A2.

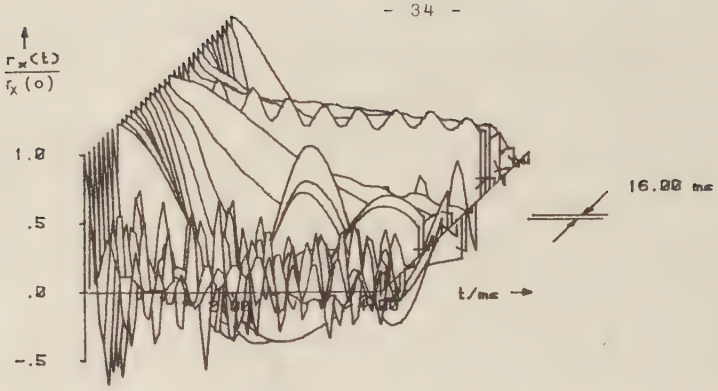


Bild 3.2 Kurzzeitautokorrelationsfunktionen(M116M)
Integrationszeit= 16.00 ms

Dabei ist $|E(z)|^2 = \sigma_e^2$ eine Konstante und H(z) ist die z-Übertragungsfunktion eines linearen, zeitinvarianten Filters.

3.1 Autoregressive Signalquelle als Modellsignalquelle für den Sprachprozeß

Aufgrund der oben angestellten Betrachtungen läßt sich folgendes einfache, lineare und für kurze Segmentlängen annähernd konstante Modell des Sprachprozesses angeben [20,21] (Bild 3.3). In diesem Modell ist e(nT) die sogenannte Anregungsfunktion. Für stimmhafte Laute kann man e(nT) durch eine periodische Impulsfolge mit der Periodenlänge n_pT (Pitch-Periode) beschreiben.

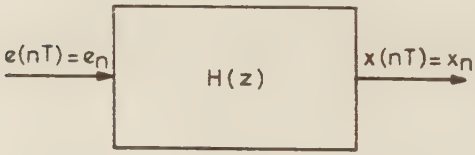


Bild 3.3 Lineares Spracherzeugungsmodell

Für stimmlose Laute sind die e(nT) Abtastwerte einer rauschartigen Zeitfunktion mit konstantem Leistungsspektrum, wobei die Zeitfunktion Musterfunktion eines Zufallsprozesses ist, der durch seine WDF beschrieben wwerden kann. Die Übertragungsfunktion H(z) ist in diesem Modell eine Funktion der Form:

$$H(z) = \frac{\cdot}{A(z)} \qquad (3.2)$$

mit A(z) nach Gleichung (2.51). H(z) gibt dabei im wesentlichen die Eigenschaften des Vokaltraktes |21| wieder. Dieses einfache und allgemein betrachtete Modell des Sprachprozesses ist also ein autoregressives Modell nach Gleichung (2.50). Für stimmlose Laute, d.h. für rauschartige Sprachsignale ist es sogar fast eine gaußsche, autoregressive Quelle. Wenn man einen solchen Prozeß nach Abschnitt 2.6 optimal codieren will, ist es zunächst einmal notwendig, die Koeffizienten a_k von A(z) zu ermitteln.

3.2 Koeffizientenbestimmung der Modellsignalquelle aus realem Sprachsignal

Die Bedeutung des linearen Modells nach Gleichung (2.50) liegt darin, daß dessen Koeffizienten direkt aus der Folge der Abtastwerte $\{x_n\}$ gewonnen werden können. Es gilt ja:

$$e_n = x_n + \hat{x}_n \tag{3.3}$$

oder $$|E(z)|^2 = |A(z)|^2 \, S_{zx}(z) \quad . \tag{3.4}$$

Dabei ist $\hat{x}_n$ die lineare Prädiktion des Abtastwertes x_n:

$$\hat{x}_n = \sum_{k=1}^{M} a_k x_{n-k} \quad . \tag{3.5}$$

Die Größe e_n ist also der Prädiktionsfehler des Filters A(z), wenn M vorangehende Abtastwerte berücksichtigt werden. Da 1/A(z) das Sprachmodell darstellt, erhält man dessen Koeffizienten auf eine bestimmte optimale Weise, wenn man bei der Schätzung eines neuen Signalwertes x_n mit dem Filter A(z) einen möglichst kleinen Prädiktionsfehler macht. Das Gütekriterium der Prädiktion soll wieder der MQF zwischen x_n und $\hat{x}_n$ für eine hinreichend große Zahl von Abtastwerten sein. Es soll also für den MQF α gelten:

$$\alpha = \sum_{\nu=n_o}^{n_1} e_\nu^2 \stackrel{!}{=} \text{Minimum} \tag{3.6}$$

oder $$\alpha = \sum_{\nu=n_o}^{n_1} \left| \sum_{k=0}^{M} a_k x_{\nu-k} \right|^2 \stackrel{!}{=} \text{Minimum} \tag{3.7}$$

mit $a_o = 1$.

Eine notwendige Bedingung, damit α minimal wird, ist:

$$\frac{\partial \alpha}{\partial a_i} = 2 \sum_{\nu=n_o}^{n_1} \left[\sum_{k=0}^{M} a_k x_{\nu-k} \right] x_{\nu-i} = 0 \quad \text{für } i=1,2,..,M \; . \quad (3.8)$$

Die Lösung dieses Minimierungsproblems hängt von der Wahl der Grenzen n_o und n_1 bzw. von der Anzahl N, der in diesem Intervall von Null verschiedenen Abtastwerte x_n ab. Es wird in jedem Fall nur ein endliches Intervall betrachtet, über das minimiert wird.

Hier soll die **Auto**korrelationsmethode (siehe z.B. |21| benutzt werden, für die sich ein schnelles rekursives Lösungsverfahren angeben läßt. Damit ist es möglich, für das Analysefilter A (z) eine andersartige Filterstruktur, das sogenannte PARCOR-Filter |21| zu bestimmen. Die in diesem Filter auftretenden Signalgrößen sind die Vorwärts- und Rückwärtsprädiktionsfehler $x^+_{n,m}$, $x^-_{n,m}$ für m = 0,1,2,...M, die bei einer Prädiktion von x_n und x_{n-m-1} mit jeweils den gleichen Abtastwerten entstehen. Die Autokorrelationsmethode der linearen Prädiktion ist nun wie folgt definiert:

$$n_o = -\infty \quad , \quad n_1 = \infty$$

und $\quad x_n = 0 \quad$ für $n < 0 \quad$ und $\quad n \geq N \quad .$

Damit ergibt sich folgendes lineare Gleichungssystem aus (3.8) zur Bestimmung der optimalen Koeffizienten a_k:

$$\sum_{k=1}^{M} a_k r_{|k-i|} = -r_i \quad \text{für } i=1,2,...,M \qquad (3.9)$$

mit $$r_l = \sum_{n=0}^{N-i-1} x_n x_{n+l} \quad \text{für } |k-i|=l=0,1,...,M \qquad (3.10)$$

Die r_l sind dabei die Autokorrelationskoeffizienten des Signalsegmentes.

3.3 PARCOR-Analysefilter zur rekursiven Lösung des linearen Gleichungssystems der Autokorrelationsmethode

Die Koeffizientenmatrix des linearen Gleichungssystems

$$\begin{bmatrix} r_o & r_1 & \cdots & r_{M-1} \\ r_1 & r_o & \cdots & r_{M-2} \\ \vdots & & & \vdots \\ r_{M-1} & r_{M-2} & \cdots & r_o \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix} = \begin{bmatrix} -r_1 \\ -r_2 \\ \vdots \\ -r_M \end{bmatrix} \qquad (3.11)$$

(3.9) besitzt Töplitzform. Zur Auflösung eines solchen Gleichungssystems wurde von Levinson |31| ein rekursiver Lösungsalgorithmus angegeben:

$$\left.\begin{aligned} a_{m+1,0} &= a_{m,0} = 1 \\ a_{m+1,1} &= a_{m,1} - k_{m+1} a_{m,m} \\ a_{m+1,2} &= a_{m,2} - k_{m+1} a_{m,m-1} \\ &\vdots \\ a_{m+1,m} &= a_{m,m} - k_{m+1} a_{m,1} \\ a_{m+1,m+1} &= k_{m+1} \end{aligned}\right\} \text{für } m=0,1,\ldots,M-1 \quad . \qquad (3.12)$$

$$\text{mit} \quad k_{m+1} = \frac{\sum_{i=0}^{m} a_{m,i} r_{m+1-i}}{\sum_{i=0}^{m} a_{m,i} r_i}$$

Dieser Lösungsalgorithmus ist besonders dadurch gekennzeichnet, daß nach jedem Rekursionsschritt ein vollständiges Analysefilter mit m+1 Stufen zur Verfügung steht. Ein m+1-stufiges Analysefilter $A_{m+1}(z)$ zur Bestimmung des minimalen Prädiktionsfehlers $e_{m+1}(nT) = x^+_{m+1}(nT)$ ergibt sich also aus einem m-stufigen Analysefilter gemäß folgender Gleichung:

$$A_{m+1}(z) = A_m(z) - k_{m+1} B_m(z) \qquad (3.13)$$

mit $$A_m(z) = \sum_{k=0}^{m} a_{m,k} z^{-k}, \quad A_o(z) = 1 \quad , \qquad (3.14)$$

$$B_m(z) = z^{-(m+1)} A_m(z^{-1}), \; B_o(z) = z^{-1} \quad . \qquad (3.15)$$

Durch Multiplikation der Werte $a_{m+1,k}$, k=0,1,...,m auf der linken Seite in (3.12) mit $z^{-(m+1-k)}$, k=0,1,...,m erhält man:

$$B_{m+1}(z) = z^{-1} \left[B_m(z) - k_{m+1} A_m(z) \right] \quad . \qquad (3.16)$$

Gleichung (3.13) und (3.16) lassen sich auch in Matrixform schreiben:

$$\begin{bmatrix} A_{m+1}(z) \\ B_{m+1}(z) \end{bmatrix} = \begin{bmatrix} 1 & -k_{m+1} \\ -z^{-1} k_{m+1} & z^{-1} \end{bmatrix} \begin{bmatrix} A_m(z) \\ B_m(z) \end{bmatrix} \quad . \qquad (3.17)$$

Das sich auf diese Weise ergebende Analysefilter $A_M(z) = A(z)$ zeigt Bild 3.4. Die Koeffizienten k_{m+1} sind die sogenannten Re-

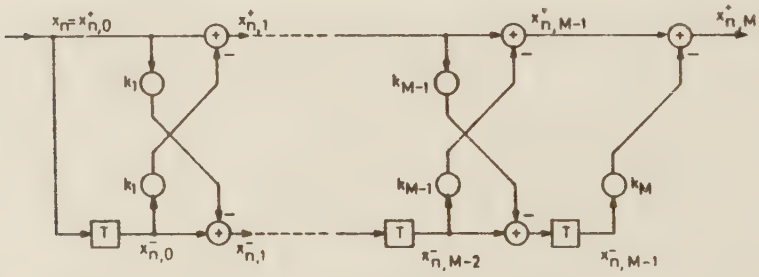


Bild 3.4 PARCOR-Analysefilter (M-Stufen)

$x^+_{n,m}$ = Vorwärtsprädiktionsfehler

$x^-_{n,m}$ = Rückwärtsprädiktionsfehler

} m = 0,1, ... ,M-1

flexionskoeffizienten. Von großer Bedeutung für die Realisierung einer fortlaufenden Sprachanalyse ist die Eigenschaft dieser Reflexionskoeffizienten, daß sie nicht nur nach Gleichung (3.12) durch die Autokorrelationswerte r_l bestimmt sind, sondern sich auch durch die Korrelation zwischen $x^+_{n,m}$ und $x^-_{n,m}$ (partielle Korrelation, PARCOR-Koeffizienten) berechnen

lassen. Es gilt ja zur Bestimmung der Filterkoeffizienten für $A_m(z)$:

$$\sum_{\nu=n_o}^{n_1} \left[\sum_{k=0}^{m} a_{m,k} x_{\nu-k} \right] x_{\nu-i} = 0 \quad i=1,2,\ldots,m \quad , \qquad (3.18)$$

da nach jedem Rekursionsschritt eine gültige Lösung für dieses Analysefilter vorliegt. Für die weitere Herleitung ist die Benutzung solcher Korrelationsausdrücke nach (3.18) zu umständlich. Es soll deshalb eine symbolische Schreibweise mit Hilfe der z-Übertragungsfunktion bestimmter Filter definiert werden. Der Ausdruck auf der linken Seite von (3.18) soll durch Bild 3.5 definiert werden:

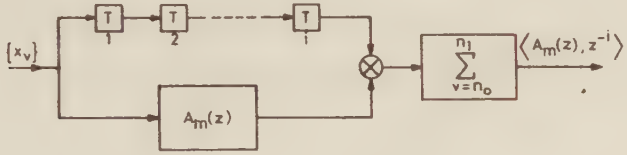


Bild 3.5 Korrelation von $x_{\nu-i}$ und $\sum_{k=0}^{m} a_{m,k} x_{\nu-k}$

Die symbolische Schreibweise für die Bedingung (3.18) (Orthogonalitätsbedingung) ist also:

$$\left\langle A_m(z), z^{-i} \right\rangle = 0 \quad , \quad \text{für } i=1,2,\ldots,m \; . \qquad (3.19)$$

Diese Bedingung muß auch für ein Filter $A_{m+n}(z)$ für i=m+n bei Beachtung von (3.13) gelten, unter der Voraussetzung, daß bei Addition solcher symbolischer Ausdrücke möglich ist:

$$\left\langle A_{m+1}(z), z^{-(m+1)} \right\rangle = \left\langle A_m(z), z^{-(m+1)} \right\rangle - k_{m+1} \left\langle B_m(z), z^{-(m+1)} \right\rangle \qquad (3.20)$$

$$\text{oder} \qquad k_{m+1} = \frac{\left\langle A_m(z), z^{-(m+1)} \right\rangle}{\left\langle B_m(z), z^{-(m+1)} \right\rangle} \; . \qquad (3.21)$$

Der Koeffizient k_{m+1} ist also, bis auf das Vorzeichen, der Quotient der Korrelation zwischen $x^{+}_{n,m}$ und x_{n-m-1} bzw. $x^{-}_{n,m}$ und x_{n-m-1} mit

$$x_{n-m-1} = x^{-}_{n,m} - \sum_{k=1}^{m} b_{m,k} x_{n-k} \; . \qquad (3.22)$$

Die $b_{m,k}$ sind die Koeffizienten von $B_m(z)$. Aus Gleichung (3.18) erkennt man, daß bei der Ausführung der jeweiligen

Korrelation in (3.21) entsprechend (3.22) nur $x^-_{n,m}$ einen Beitrag liefert, der ungleich Null ist. Aus (3.21) wird damit:

$$k_{m+1} = \frac{\left\langle A_m(z), B_m(z) \right\rangle}{\left\langle B_m(z), B_m(z) \right\rangle} \tag{3.23}$$

oder

$$k_{m+1} = \frac{\sum_{\nu=n_o}^{n_1} x^+_{\nu,m} x^-_{\nu,m}}{\sum_{\nu=n_o}^{n_1} |x^-_{\nu,m}|^2} \quad . \tag{3.24}$$

Gleichzeitig läßt sich bei der Autokorrelationsmethode zeigen, daß auch gilt:

$$k_{m+1} = \frac{\sum_{\nu=n_o}^{n_1} x^+_{\nu,m} x^-_{\nu,m}}{\sum_{\nu=n_o}^{n_1} |x^+_{\nu,m}|^2} \quad . \tag{3.25}$$

3.4 PARCOR-Synthesefilter

Mit Hilfe der Reflexionskoeffizienten k_{m+1} läßt sich nun entsprechend eine PARCOR-Synthesefilterstruktur angeben, mit der das Signal x_n aus dem Fehlersignal $e_n = x^+_{n,M}$ ohne den Umweg über die Prädiktorkoeffizienten $a_{M,k}$ erzeugt werden kann. Durch Multiplikation des Gleichungssystems (3.17) mit $E(z)/A(z)$ und Umstellen der ersten Gleichung erhält man:

$$\left.\begin{aligned} \frac{A_m(z)E(z)}{A(z)} &= \frac{A_{m+1}(z)E(z)}{A(z)} - k_{m+1}\frac{B_m(z)E(z)}{A(z)} \\ \frac{zB_{m+1}(z)E(z)}{A(z)} &= -k_{m+1}\frac{A_m(z)E(z)}{A(z)} + \frac{B_m(z)E(z)}{A(z)} \end{aligned}\right\} \text{für } m=M-1,M-2,\ldots,0 \; . \tag{3.26}$$

Es gilt insbesondere:

$$\frac{A_o(z)E(z)}{A(z)} = X(z) \quad , \quad A_o(z) = 1 \quad , \tag{3.27}$$

$$\frac{A_M(z)E(z)}{A(z)} = E(z) \quad , \quad A_M(z) = A(z) \quad . \qquad (3.28)$$

X(z) ist dabei die z-Transformierte von $\{x_n\}$. Bild 3.6 zeigt die Struktur dieses Synthesefilters.

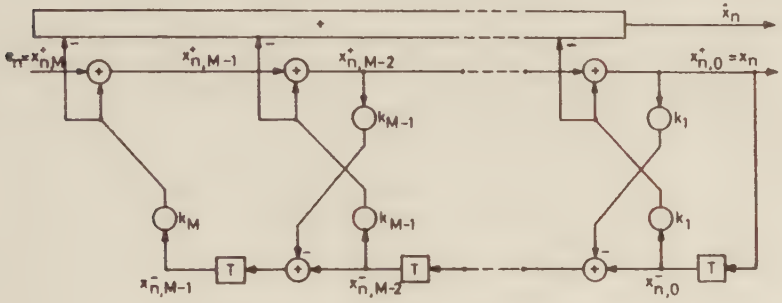


Bild 3.6 PARCOR - Synthesefilter mit Prädiktion

Ein solches PARCOR-Synthesefilter ist ein Wellendigitalfilter |32|, wobei die $x^+_{n,m}$ und $x^-_{n,m}$ die Wellengrößen darstellen. Die Koeffizienten k_{m+1} und das Fehlersignal e_n müssen durch eine entsprechende Analyse der Folge $\{x_n\}$ ermittelt werden.

Die Filterstruktur nach Bild 3.6 enthält zusätzlich noch die Berechnung der Prädiktion $-\hat{x}_n$ des Abtastwertes x_n, so wie es später, in dem in dieser Arbeit betrachteten Deltamodulationssystem, notwendig ist. Der Abtastwert x_n ergibt sich aufgrund folgender Beziehung:

$$x_n = \sum_{m=1}^{M} x^-_{n,m-1} k_m + e_n \quad , \qquad (3.29)$$

d. h.,

$$\hat{x}_n = -\sum_{m=1}^{M} x^-_{n,m-1} k_m \quad . \qquad (3.30)$$

Die einzelnen Terme in (3.30) können abgespeichert und auch noch zur Synthese benutzt werden, falls $\hat{x}_n$ vollständig vor dem Eintreffen eines neuen Wertes e_n zur Verfügung stehen soll.

Die Stabilität des Synthesefilters 1/A(z) ist dann gewährleistet, wenn alle Nullstellen des Polynoms A(z) innerhalb des Einheitskreises liegen. Das ist aber gleichbedeutend mit der Bedingung |21|:

$$|k_m| < 1 \quad \text{für} \quad m = M, M-1, \ldots, 1 \quad . \tag{3.31}$$

Das Problem der Beachtung dieser Stabilitätsbedingung besteht darin, daß sie bei der Berechnung von k_m durch (3.23) oder (3.24) nicht notwendigerweise erfüllt ist, wenn die Berechnung mit endlicher Signalwortlänge durchgeführt wird.

Es bietet sich nach |33| auch an, k_m als den in der Statistik definierten Kreuzkorrelationskoeffizienten zwischen $x^+_{n,m}$ und $x^-_{n,m}$ aufzufassen:

$$k_{m+1} = \frac{\sum_{\nu=n_o}^{n_1} x^+_{\nu,m}\, x^-_{\nu,m}}{\left[\sum_{\nu=n_o}^{n_1} x^{+\,2}_{\nu,m} \cdot \sum_{\nu=n_o}^{n_1} x^{-\,2}_{\nu,m} \right]^{\frac{1}{2}}} \tag{3.32}$$

Der Aufwand bei der Berechnung ist allerdings größer. Im folgenden wird aber noch angegeben, welche Vereinfachungen man einführen kann, um zu vernünftigen und nicht zu aufwendigen Verfahren zur Schätzung von k_m zu gelangen.

4. DM mit adaptivem Rekonstruktionsnetzwerk

Bei der in dieser Arbeit beabsichtigten Optimierung der DM für Sprachsignale, so wie sie in Abschnitt 3 beschrieben wurden, soll das einfache Systemkonzept der DM soweit wie möglich beibehalten werden. Ausgangspunkt für die Optimierung sind die Modellbeschreibung des Sprachprozesses und die Forderungen der optimalen Baumcodierung für solche Prozesse nach Abschnitt 2.6.2. Da die Parameter des Sprachmodells sich zeitlich ändern, muß sich auch das Rekonstruktionsnetzwerk des Deltamodulators zeitlich ändern. Alle anderen Systemteile sollen dagegen erhalten bleiben.

Ein solches Deltamodulationssystem hat demnach prinzipiell das Aussehen nach Bild 4.1. Als adaptives Rekonstruktions-

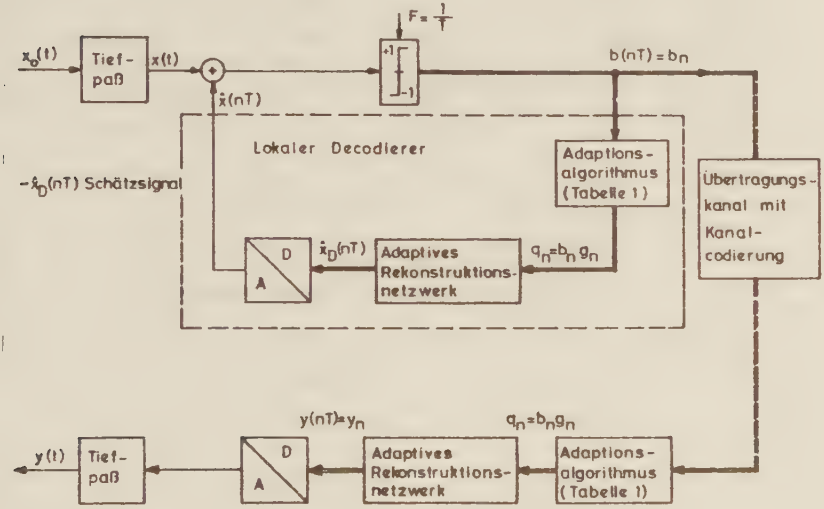


Bild 4.1 Deltamodulationssystem mit adaptivem Rekonstruktionsnetzwerk.

netzwerk wird ein PARCOR-Synthesefilter mit Prädiktion nach Bild 3.6 verwendet. Eine genauere Begründung für die Verwendung dieses Netzwerkes wird in den folgenden Abschnitten (4.1 und 4.3) gegeben. Insbesondere ist das PARCOR-Synthesefilter mit Prädiktion in Bild 4.6 dargestellt. Der Algorithmus zur

Adaption der Reflexionskoeffizienten wird in Abschnitt 4.4 erläutert. Aufgrund von Rechnersimulationen und Stabilitätsbetrachtungen (Abschnitt 5) für ein solches Deltamodulationssystem wird angegeben, welche Anfangswerte die Reflexionskoeffizienten besitzen müssen und aus welchen Wertebereichen diese Koeffizienten Werte annehmen dürfen, damit ein stabiler Deltamodulationsprozeß gewährleistet ist.

Diese Art der DM wird im folgenden DM mit adaptivem PARCOR-Prädiktionsfilter (APPDM) genannt. Die Kurzbezeichung für die jeweils verwendeten Parameter wird im Anhang A4 erläutert.

Bei der Optimierung dieses Systems wird zunächst nur das adaptive Rekonstruktionsnetz betrachtet. Das Eingangssignal q_n dieses Netzwerkes ist die mit einem bit quantisierte Differenz zwischen x(t) und dem Schätzsignal $-\hat{x}(nT)$ (siehe auch Gleichung (3.3)). Der Algorithmus zur Adaption der Quantisierungsstufe g_n (Tabelle 1) wird aus [34,35] übernommen. Da

Tabelle 1 Adaptionsalgorithmus

$g_{min} = g_{min} k^0, g_{max} = g_{min} k^{i_{max}}$

$g_n = g_{min} k^i,\ i = 0,1,\ldots,\ i_{max}$

Bitmuster der Länge N	Adaptionsfaktor k>1
$b_n\ b_{n-1}\ b_{n-2} \cdots b_{n-N+1}$	$0 < g_{min} \le g_n \le g_{max}$
1 1 1 .... 1 -1 -1 -1 .... -1	$g_n = g_{n-1} k$
1 -1 1 .... $-1^{(N-1)}$ -1 1 -1 .... -1^N	$g_n = g_{n-1} / k$
sonst	$g_n = g_{n-1}$

dort die DM bei relativ hohen Abtastraten (>32 kHz) betrachtet wurde, muß dieser Algorithmus nicht unbedingt optimal sein. Ob und wieweit Änderungen an diesem Algorithmus vorgenommen werden müssen , soll aber bei der Optimierung des adaptiven Rekonstruktionsnetzwerkes mit untersucht werden. Ähnliche Algorithmen zur Quantisierungsstufenadaption werden z.B. auch in [10] angegeben.

Die Anforderungen an ein solches Deltamodulationssystem und die sich daraus ergebenden Vorteile gegenüber anderen digitalen Modulationsverfahren sollen hier noch einmal kurz zusammengestellt werden:

1. Zu jedem Abtastzeitpunkt wird nur ein Binärzeichen b_n erzeugt, das zum Empfänger übertragen wird (keine Wortsynchronisation).

2. Außer dieser Binärzeichenbestimmung ist keine weitere A/D-Wandlung notwendig.

3. Bei einer solchen "A/D-Wandlung" wird kein Abtast-Halteglied benötigt.

4. Außer der Übertragung von b_n soll keine zusätzliche Information über die augenblickliche Quantisierungsstufe und die Koeffizienten des adaptiven Rekonstruktionsnetzwerkes erfolgen. Hierbei ist allerdings vorausgesetzt, daß zu Anfang der Übertragung der Gleichlauf zwischen Sender und Empfänger auf geeignete Weise sichergestellt wird und daß während der Übertragung keine Fehler auftreten, die diesen Gleichlauf stören könnten.

5. Die Abtastrate ist im allgemeinen größer als die Nyquistrate. Damit wird das Tiefpaßfilter gegenüber PCM- und DPCM-Systemen einfacher.

4.1 Vergleich von DM und optimaler Baumcodierung

Ein Deltamodulator verwirklicht die einfachste denkbare Baumcodierung. Der Codebaum hat das Aussehen nach Bild 4.2.

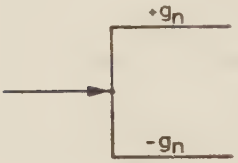


Bild 4.2 Codebaum der DM

Die Zweige des Codebaums sind mit den Werten (Symbolen) $+g_n$ und $-g_n$ besetzt. Die Informationsrate dieses Codes ist:

$$R_{CB} = \frac{1}{\beta} \ld\alpha = 1 \text{ bit/Symbol.} \qquad (4.1)$$

Die Blocklänge ist N=1. Die Auswahl des optimalen Pfades, der den minimalen MQF liefert, erfolgt aufgrund der Differenz zwischen Signalwert x(t)=x(nT) und Prädikation $-\hat{x}(nT)$:

$$|x(nT) - y(nT)| = \min_{q_n \in \{-g_n, g_n\}} |x(nT)+\hat{x}(nT)+q_n| \quad . \qquad (4.2)$$

Bei der optimalen Baumcodierung ist es aber zunächst notwendig, daß die Eingangssignalfolge $\{v_n\}$(siehe Bild 2.7) des Rekonstruktionsnetzwerkes statistisch unabhängig ist, damit man die Informationsrate des Codes aus der Baumstruktur alleine bestimmen kann. Das ist aber bei der DM nicht der Fall. Die Suche des optimalen Pfades ist abhängig von $x(nT)+\hat{x}(nT)$ und damit von vorhergehenden Entscheidungen. Die Eingangsfolge $\{q_n\}$ des Rekonstruktionsnetzwerkes zur Codeerzeugung ist also prinzipiell eine statistisch abhängige Folge. Damit ist die effektive Informationsrate R_{DM} des Deltamodulatorcodes kleiner als die Informationsrate des Codebaums nach Bild 4.2. Es gilt also:

$$R_{DM} < R_{CB} = 1 \text{ bit/Symbol} \quad . \qquad (4.3)$$

Eine optimale Baumcodierung setzt zusätzlich eine hinreichend große Blocklänge voraus. Das ist aber bei der DM mit N=1 ebenfalls als nicht gegeben anzusehen.

Aus diesen beiden Gründen ist es nicht möglich, daß ein Deltamodulationssystem die nach der RDT möglichen SNR-Werte für die zur Verfügung stehende Informationsrate des Codebaums

erreicht.

Weiter ist zur Erzeugung eines optimalen Codes notwendig, daß das Rekonstruktionsnetzwerk aufgrund der Ergebnisse von Abschnitt 2.6.2 an die Signalquelle angepaßt wird. Es soll einmal davon abgesehen werden, daß die Voraussetzungen, die zur Bestimmung des optimalen Rekonstruktionsnetzwerkes führten, durch die oben genannten Gründe nicht mehr genau gegeben sind.

Wenn man Sprachsignale mit sich ändernden statistischen Eigenschaften optimal codieren will, muß man zur Optimierung für jedes Sprachsegment folgende Schritte durchführen:

1. Berechnung der Kovarianzmatrix $\underline{R}_x$ für das zu codierende Sprachsegment.

2. Bestimmung der Eigenwerte von $\underline{R}_x$ und Berechnung des minimal möglichen MQF's D nach der RDT für die gegebene Kanalkapazität unter Berücksichtigung der vorgesehenen Abtastrate (wenn man zur Berechnung der RDF das Gleichungspaar (2.41) und (2.42) benutzen will).

3. Danach müssen zwei Fälle unterschieden werden (siehe auch Satz 2.2):

   3.1 Es gilt $D \leq \delta$. Dann ist zur Berechnung des optimalen Rekonstruktionsnetzwerkes ein nichtlineares Gleichungssystem (2.68) zu lösen.

   3.2 Es gilt $D > \delta$. Dann kann das optimale Rekonstruktionsnetzwerk nur noch durch ein Approximationsverfahren berechnet werden.

4. Es muß ein optimaler Pfad durch den Codebaum gefunden werden. Zur Baumcodierung sind allerdings schon effiziente Suchalgorithmen bekannt [19,36].

Es ist dabei zu beachten, daß bei den in dieser Arbeit betrachteten niedrigen Übertragungsbitraten meistens der Fall 3.2 eintritt. Die Berechnung des optimalen Rekonstruktions-

netzwerks ist also sehr aufwendig. Dabei sind Verfahren zur optimalen Segmentierung von Sprachsignalen noch nicht einmal angesprochen worden. Die Anpassung des Rekonstruktionsnetzwerkes in einem Deltamodulator sollte also auf einfachere Weise erfolgen.

Wenn man das rekonstruierte Signal des Deltamodulators auf die gleiche Weise analysiert wie das zu codierende Sprachsignal, z.B. mit einem PARCOR-Filter nach Abschnitt 3.3, dann erhält man Reflexionskoeffizienten für ein entsprechendes PARCOR-Synthesefilter. Diese Reflexionskoeffizienten unterscheiden sich aber von den Reflexionskoeffizienten des zu codierenden Sprachsignals. Nach der RDT ist das rekonstruierte Signal immer mit einem Rauschen behaftet, welches mindestens den MQF D besitzt. Das bedeutet aber, daß sich bei der Analyse des rekonstruierten Signals immer ein Fehlersignal e_n ergibt, welches mindestens die Leistung $\sigma_e^2 + D$ besitzt. Dabei wird davon ausgegangen, daß das Sprachsignal durch das lineare Modell nach Bild 3.3 erzeugt worden ist und σ_e^2 die Leistung des Anregungssignals e_n ist. Bei der Analyse eines so erzeugten Sprachsignals x_n erhält man gerade das Fehlersignal (Anregungssignal) e_n mit mindestens der Leistung σ_e^2. Man kann sich die Erzeugung des rekonstruierten Signals y_n eines Deltamodulators also entsprechend Bild 4.3 vorstellen. Dabei ist σ_q^2 die Leistung des Eingangssignals q_n des Rekonstruktionsnetzwerkes. Wenn während des Deltamodulationsprozesses das rekonstruierte Signal y_n dem Eingangssignal x_n folgen kann, dann besitzen beide Signale in etwa die gleiche Leistung. Aus diesem Grunde muß dann die Verstärkung des Rekonstruktionsnetzwerkes im Deltamodulator kleiner sein als die Verstärkung des Synthesefilters H(z) nach Bild 3.3, welches das Sprachsignal x_n erzeugt.

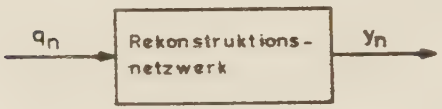


Bild 4.3 Deltamodulatorrekonstruktionsnetzwerk mit $\sigma_q^2 \geq \sigma_e^2 + D$

Betrachtet man dazu im Vergleich das optimale Leistungsspektrum $S_{zy,Opt}$ nach (2.67) für $D \leq \delta$, so erkennt man, daß die

Filterkoeffizienten des optimalen Rekonstruktionsnetzwerkes ebenfalls eine kleinere Verstärkung liefern als das Synthesefilter, welches das Sprachsignal x_n erzeugt.

Aufgrund dieser Überlegungen scheint es vernünftig zu sein, zur Berechnung der Reflexionskoeffizienten des Rekonstruktionsnetzwerkes im Deltamodulator die Signale des PARCOR-Synthesefilters (siehe auch Abschnitt 4.3) heranzuziehen, die den Vorwärts- und Rückwärtsprädiktionsfehlern eines PARCOR-Analysefilters entsprechen. Diese Signale stehen darüber hinaus auch dem Empfänger zur Verfügung. Welche SNR-Werte sich dabei, auch im Vergleich zu anderen bekannten Quellcodierungsverfahren, ergeben, kann dann nur durch Messung und Berechnung bzw. bei realen Sprachsignalen durch subjektive Hörtests ermittelt werden.

4.2 DM bei niedrigen Abtastraten

Wegen der in Abschnitt 2.5 gezeigten Abhängigkeit der RDF von der Abtastrate, ist es sinnvoll, diese Abtastrate bei einem Quellcodierer so niedrig wie möglich zu wählen. Bei fest vorgegebener Kanalkapazität werden die zu erreichenden SNR-Werte allerdings immer kleiner. Für Kanalkapazitäten < 20 kbit/s werden die erreichbaren SNR-Werte für Quellcodierungsverfahren, die bei der Nyquistrate arbeiten, und für die DM annähernd gleich. Für eine gaußsche Signalquelle mit

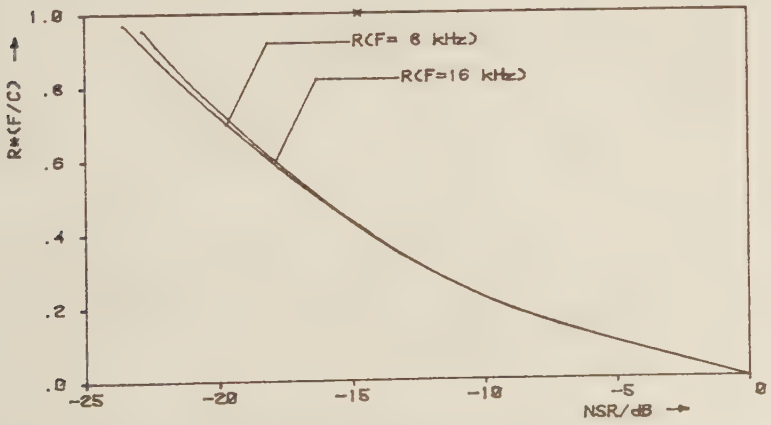


Bild 4.4 R(F,NSR)=(F/C)/Kanalkapazitaet C=16 kbit/s
Gedaechtnislaenge=1.1 ms

dem Leistungsspektrum† nach Bild 2.4 ergibt sich für eine Kanalkapazität von 16 kbit/s der Verlauf der RDF nach Bild 4.4. Die erreichbaren SNR-Werte bei F = 8 kHz und F = 16 kHz sind annähernd gleich. Ähnliche Verhältnisse ergeben sich auch für die zu erreichenden SNR-Werte bei stimmlosen Lauten (Bild 4.5).

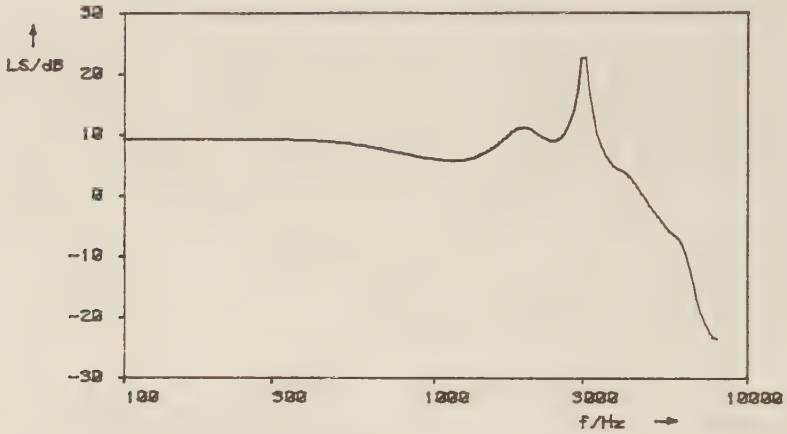


Bild 4.5a Log. Leistungsspektrum LS(f)
PARCOR-Analyse von M116M
12 Filterstufen/Integrationszeit= 16 ms
Laut:/sch/

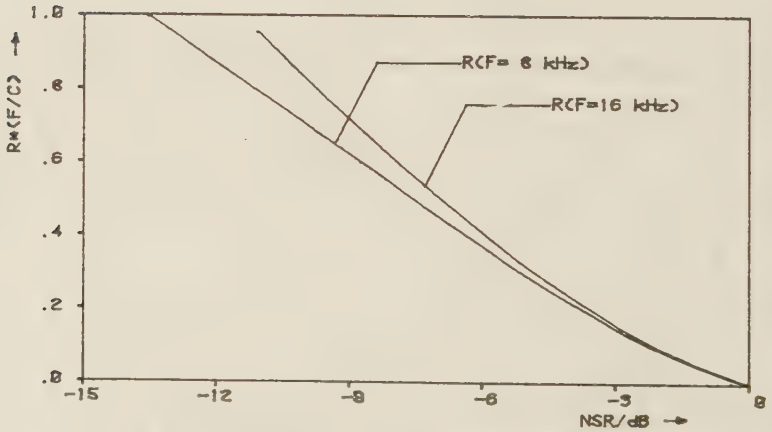


Bild 4.5b R(F,NSR)=(F/C)/Kanalkapazitaet C=16 kbit/s
Gedaechtnislaenge= .8 ms

†Kurzzeitleistugsspektrum für den Vokal |ä| in "Das Herz Münchens schlägt ..." (Anhang A2)

Hier ist der maximal mögliche SNR-Wert bei F = 8kHz aber ca. 2 dB größer als für F = 16 kHz.

Zusätzlich ist in Bild 4.4 der SNR-Wert eingezeichnet (x), der für die betrachtete gaußsche Signalquelle von einem krümmungsadaptivem Deltamodulator† erreicht wird. Dabei wurde angenommen, daß die Informationsrate dieses Deltamodulators 1 bit/Symbol beträgt. Es ist zu erkennen, daß noch ein mehr oder weniger großer Unterschied zum maximal erreichbaren SNR-Wert besteht. Dabei ist zu beachten, daß eine gewisse Parameteroptimierung für das Leistungsspektrum nach Bild 2.4 für diesen Deltamodulator vorgenommen wurde. Der erreichte SNR-Wert ist also unter Umständen bei einer umfassenden Optimierung für reale Sprachsignale für dieses spezielle Testsignal zu gut. Für die gaußsche Signalquelle mit dem Leistungsspektrum nach Bild 4.5a wurde sogar eine Geräuschleistung erzielt, die größer war als die Signalleistung.

4.3 Struktur des adaptiven Rekonstruktionsnetzwerkes

Aufgrund der Überlegungen in Abschnitt 4.1 soll eine Adaption des Rekonstruktionsnetzwerkes mit Hilfe der entsprechenden Signale dieses Netzwerkes selber durchgeführt werden. In einem PARCOR -Synthesefilter stehen die gleichen Signale zur fortlaufendne Berechnung der Reflexionskoeffizientne wie in einem PARCOR-Analysefilter zur Verfügung. Die Entscheidungen des Deltamodulators müssen nur gewährleisten, daß das rekonstruierte Signal y_n dem Eingangssignal x_n entsprechend gut folgen kann. Eine Berechnung des benötigten Schätzwertes $-\hat{x}_n$ ist nach Abschnitt 3.4 ebenfalls ohne großen zusätzlichen Aufwand mit dieser Struktur möglich. Das adaptive Rekonstruktionsnetzwerk in Bild 4.6 ist also entsprechend Bild 3.6 aufgebaut. Dabei sind e_n durch q_n, x_n durch y_n, $x^+_{n,m}$ durch $y^+_{n,m}$ und $x^-_{n,m}$ durch $y^-_{n,m}$ zu ersetzen. Der Schätzwert $-\hat{x}_n$ soll hier genauso bezeichnet werden, da auch hier das Originalsignal x_n geschätzt werden soll. Zusätzlich zur Erzeugung des rekonstruierten Signals y_n und des Schätzsignals $-\hat{x}_n$ ist jetzt noch

†Siehe Anhang A3.

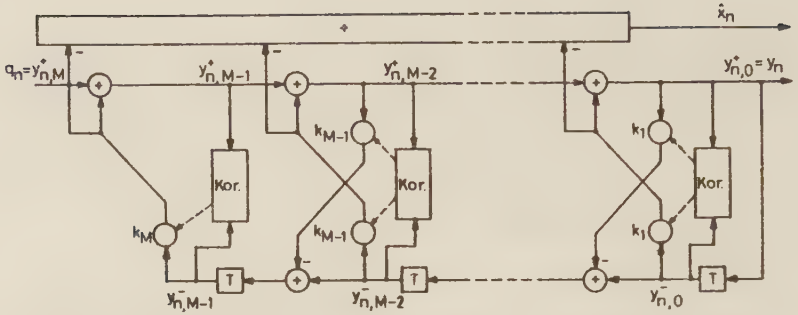


Bild 4.6 PARCOR -Synthesefilter mit Prädiktion und Berechnung der Reflexionskoeffizienten (Kor.)

die Berechnung der Reflexionskoeffizienten notwendig. Dazu sind in der Literatur |33,37,38| eine Reihe von Verfahren vorgeschlagen worden. Man kann eine Einteilung dieser Verfahren danach vornehmen, daß die Reflexionskoeffizienten nur für aufeinanderfolgende Sprachsegmente oder aber zu jedem Abtastzeitpunkt neu berechnet werden.

Bei der DM kommt nur die zweite Möglichkeit in Betracht, da sonst die Segmentgrenzen zusätzlich synchronisiert werden müßten. Selbst bei den in dieser Arbeit betrachteten niedrigen Abtastraten bedeutet diese Forderung allerdings, daß sehr schnelle und effektive Verfahren zur Berechnung der Reflexionskoeffizienten notwendig sind, besonders wenn hohe Synthesefiltergrade zu realisieren sind.

Bei M Stufen des Synthesefilters sind 2M Korrelationsausdrücke zu berechnen. Nach M Divisionen stehen dann M neue Reflexionskoeffizienten zur Verfügung. Wenn man die Reflexionskoeffizienten mit Hilfe der Autokorrelationswerte des rekonstruierten Signals und des rekursiven Algorithmus (3.12) berechnet, dann sind nur M+1 Korrelationsausdrücke zu bestimmen. Zusätzlich ist aber der Rechenalgorithmus durchzuführen, der ebenfalls eine bestimmte Rechenzeit und Rechengenauigkeit erfordert. Wenn solche Algorithmen mit beschränkter Stellenzahl der Rechenregister ablaufen, dann ist aber eine genaue Skalierung der betrachteten Rechengrößen notwendig |39|.

Die Rechenalgorithmen zur Berechnung der Reflexionskoeffizienten, die sich aus der Struktur des PARCOR-Filters ergeben,

sind also einfacher als der Umweg über die Autokorrelationswerte und den erwähnten Lösungsalgorithmus. Zusätzlich ergibt sich eine einfache Möglichkeit zur Stabilitätsprüfung des PARCOR-Synthesefilters.

4.4 Vereinfachte und schnelle Verfahren zur Bestimmung der Reflexionskoeffizienten

Eine Vereinfachung der Berechnung von k_{m+1} ergibt sich, wenn man im Nenner von (3.32) statt des geometrischen Mittels (k^g_{m+1}) das arithmetische Mittel (k^a_{m+1}) von

$$\sum_{\nu=n_0}^{n_1} x^{+2}_{\nu,m} \qquad \text{und} \qquad \sum_{\nu=n_0}^{n_1} x^{-2}_{\nu,m}$$

benutzt [33]:

$$k^a_{m+1} = \frac{2 \sum_{\nu=n_0}^{n_1} x^{+}_{\nu,m}\, x^{-}_{\nu,m}}{\sum_{\nu=n_0}^{n_1} x^{+2}_{\nu,m} + \sum_{\nu=n_0}^{n_1} x^{-2}_{\nu,m}} \quad . \tag{4.4}$$

Wegen der bekannten Beziehung zwischen geometrischem und arithmetischem Mittel gilt:

$$|k^a_{m+1}| \leq |k^g_{m+1}| \quad . \tag{4.5}$$

Gleichung (4.4) liefert also ebenfalls Reflexionskoeffizienten für ein stabiles Synthesefilter. Dabei muß aber keine Wurzel gezogen werden.

In einem Deltamodulator ist keine Verzögerung des Signals über mehrere Abtastwerte vorgesehen, so daß eine Berechnung der Reflexionskoeffizienten immer auf der Grundlage schon vergangener Abtastwerte des rekonstruierten Signals erfolgen muß. Daher ist es wünschenswert, daß die aktuellen Abtastwerte stärker gewichtet werden als die vergangenen. Dazu wird eine Gewichtsfunktion

$$w_n = w(nT) \tag{4.6}$$

eingeführt. Damit erhält man die neue Berechnungsvorschrift für k^a_{m+1} zum Zeitpunkt n:

$$k^a_{m+1} = \frac{2 \sum_{\nu=-\infty}^{n} w_{n-\nu}\, x^+_{\nu,m} x^-_{\nu,m}}{\sum_{\nu=-\infty}^{n} w_{n-\nu}\, (x^{+2}_{\nu,m} + x^{-2}_{\nu,m})} \tag{4.7}$$

Diese Gewichtung kann man mit einer rekursiven oder nichtrekursiven Filterung der entsprechenden Signale erreichen. Ein nichtrekursives Filter ist für die Realisierung eines schnellen Gewichtungsalgorithmus ungeeignet, wenn man Gewichtsfunktionen realisieren will, die nicht konstant sind für eine bestimmte Anzahl von Abtastwerten und dann verschwinden. Bei Sprachsignalen liegen günstige Werte für die Länge dieses Intervalls von Abtastwerten zwischen 10 ms und 20 ms. Bei einer Abtastrate von 16 kHz bedeutet das aber schon ca. 2M•256 Speicherplätze (entsprechend 16 ms) für die Wellengrößen des Synthesefilters. Bei einer Gewichtsfunktion, die nicht konstant ist, müßten alle Speicherinhalte zu jedem Abtastzeitpunkt mit einem unterschiedlichen Gewichtsfaktor multipliziert werden.

Es kommen nur Gewichtsfunktionen in Frage mit der Eigenschaft:

$$\begin{aligned} w_n &\geq 0 \qquad \text{für} \quad n \geq 0\,, \\ w_n &= 0 \qquad \text{für} \quad n < 0. \end{aligned} \tag{4.8}$$

Die Gewichtsfunktion w_n darf nur positive Werte annehmen, da hier Leistungs- bzw. Korrelationswerte und nicht Signalwerte gewichtet werden. Bei negativen Werten der Gewichtsfunktion würden sich sonst bestimmte Anteile aufheben und die Gewichtung wäre nicht sinnvoll. Insbesondere dürfen bei der gewichteten Summierung von Leistungswerten keine negativen Summenglieder auftreten.

Die Gewichtung der Summenglieder kann entsprechend der Impulsantwort eines rekursiven Filters erfolgen. Die einfachste

Impulsantwort eines stabilen und kausalen rekursiven Filters ist:

$$w_n = c^n \qquad \text{für} \qquad n \geq 0 \quad ,$$

$$w_n = 0 \qquad \text{für} \qquad n < 0 \qquad (4.9)$$

mit $0 \leq c < 1$.

Die z-Übertragungsfunktion W(z) eines solchen Filters mit $z = e^{j\varphi}$ und $W(\varphi=0)$ lautet:

$$W(z) = \frac{1 - c}{1 - cz^{-1}} \quad . \qquad (4.10)$$

Die Zeitkonstante des Gewichtungsintervalls (Integrationszeit) dieses Filters beträgt:

$$T_i = - \frac{T}{\ln c} \quad . \qquad (4.11)$$

Der Koeffizient c kann so gewählt werden, daß bei der digitalen Realisierung von W(z) keine echten Multiplikationen auftreten. Nach Itakura und Saito [33] kann man dann einen Algorithmus zur Auswertung der Gleichung 4.7 angeben, der in Bild 4.7 dargestellt ist. Abgesehen von Additionen und Vorzeichenumkehr treten nur zwei echte Multiplikationen und eine Division auf.

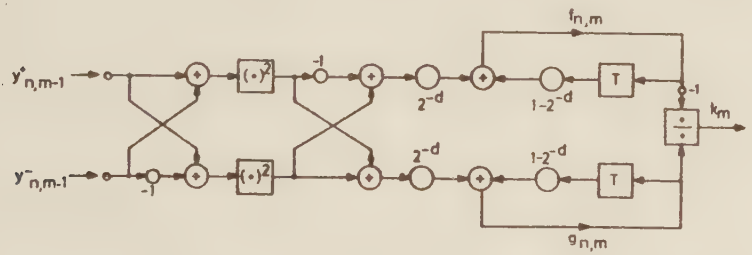


Bild 4.7 Berechnungsalgorithmus (Kor.) für
k_m, m = 1,2,..., M;
$c = 1 - 2^{-d}$; d > 0; d ganzzahlig

Zusätzlich ist noch zu sagen, daß bei den Multiplikationen nur die Beträge[†] von $(y^{+}_{n,m} + y^{-}_{n,m})$ und $(y^{+}_{n,m} - y^{-}_{n,m})$ multipliziert werden müssen. Zur Darstellung der Reflexionskoeffizienten sind unter Umständen nur 3 bis 7 bit notwendig|21,39|. Dadurch lassen sich die Division aber auch die beiden Multiplikationen pro Stufe des PARCOR-Synthesefilters entsprechend schnell durchführen.

[†] Statt $x^{+}_{n,m}$ und $x^{-}_{n,m}$ werden hier $y^{+}_{n,m}$ und $y^{-}_{n,m}$ benutzt, da ja die Reflexionskoeffizienten im Deltamodulator aus den entsprechenden Signalen im Deltamodulator berechnet werden sollen.

5. Stabilität des Deltamodulationsprozesses

In der Literatur|40| ist die Stabilitätsuntersuchung eines Deltamodulators mit einfacher Integration (ideal und nichtideal) mit fester Quantisierungsstufe beschrieben. Die Untersuchung wurde für ein stationäres Zufallssignal mit vorgegebenem Leistungsspektrum durchgeführt. Es wurden dabei Grenzen für die Varianz des Quantisierungsgeräusches angegeben. Die Sicherstellung einer solchen Grenze reicht aber in der Praxis nicht aus, um eine geräuscharme Nachrichtenübertragung zu gewährleisten. Die erwähnte Stabilitätsuntersuchung ist statistischer Natur. Es ist aber bekannt |10,41,42|, daß bei der DM z.B. Leerlaufgeräusche auftreten können, die die Übertragung erheblich stören oder sogar unmöglich machen können. Die Untersuchung solcher Störungen soll Gegenstand dieses Kapitels sein.

In |10, 41,42| wurden solche Untersuchungen für Deltamodulatoren mit festem Rekonstruktionsnetzwerk und Abtastraten, die sehr viel größer als die Nyquistrate waren, durchgeführt. In der vorliegenden Arbeit, bei zeitveränderlichem Rekonstruktionsnetzwerk, ist diese Untersuchung erheblich aufwendiger. Es werden Bedingungen für die Existenz dieser Störungen angegeben. Daran schließt sich eine Betrachtung über die Anregung solcher Störungen auf der Grundlage der RDT an.

5.1 Grenzzyklusschwingungen (Grenzzyklen) in nichtlinearen Abtastsystemen

Ein Deltamodulator ist ein nichtlineares, rückgekoppeltes Abtastsystem (Bild 5.1). In einem solchen nichtlinearen System können Grenzzyklen auftreten |27|. Bei dieser Untersuchung ist es nicht wichtig ob

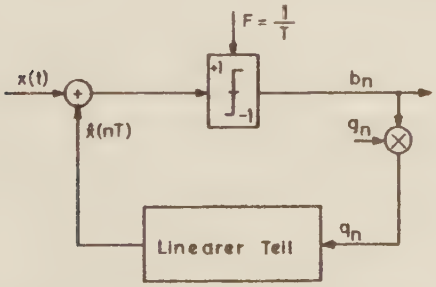


Bild 5.1 Nichtlineares rückgekoppeltes Abtastsystem

eine einmal angeregte Grenzzyklusschwingung nach kurzer Zeit von selbst wieder abklingt oder ob sie bestehen bleibt. Wenn Grenzzyklen auftreten, deren Grundfrequenz im Übertragungsbereich liegt, so wird die Übertragung mehr oder weniger gestört. Es ist also vor allem wichtig, die Existenz solcher Grenzzyklen oder aber deren Anregung (mit großer Sicherheit) auszuschließen.

Unter einem Grenzzyklus in einem Abtastsystem versteht man eine Folge von Abtastwerten an irgendeiner Stelle im System, die sich periodisch wiederholen. Dabei ist das Eingangssignal x(t) eine periodische Funktion der Zeit. Bei zufälligem Eingangssignal x(t) bleiben die angeregten Grenzzyklen im allgemeinen nicht erhalten. Wichtige in diesem Zusammenhang zu betrachtende Sonderfälle sind:

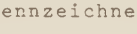
$$x(t) = 0 \qquad \forall\, t$$

und

$$x(t) = \text{const.} \neq 0 \qquad \forall\, t \quad .$$

In einem Deltamodulator bietet es sich zur Charakterisierung einer bestimmten Grenzzyklusschwingung, kurz Mode genannt, an, die binäre Ausgangssignalfolge $\{b_n\}$ heranzuziehen. Die b_n sollen in diesem Fall Werte aus der Menge der binären Zeichen {0,1} annehmen können. Die Aufeinanderfolge dieser binären Zeichen mit der Periodenlänge L soll einen bestimmten Mode kennzeichnen:

z.B. ...001011001011001011... . (5.1)



Die Folge dieser Zeichen beschreibt einen Mode mit der Periodenlänge L=6. Eine kürzere Schreibweise erhält man, wenn man einen Mode durch die kleinste Oktalzahl, die alle binären Zeichen einer Periode berücksichtigt und die Periodenlänge L, kennzeichnet. Für das Beispiel (5.1) sieht diese Schreibweise folgendermaßen aus:

5/13 - Mode (5.2

5.2 Existenzbedingung für beliebig vorgebbare Moden bei einem Deltamodulator mit M-stufigem PARCOR-Synthesefilter und adaptiver Quantisierungsstufenbestimmung bei Eingangssignal $x(t) \equiv 0$

Grundlage für die Bestimmung der Existenz eines bestimmten Modes ist der Deltamodulator nach Bild 5.2. Das Rekon-

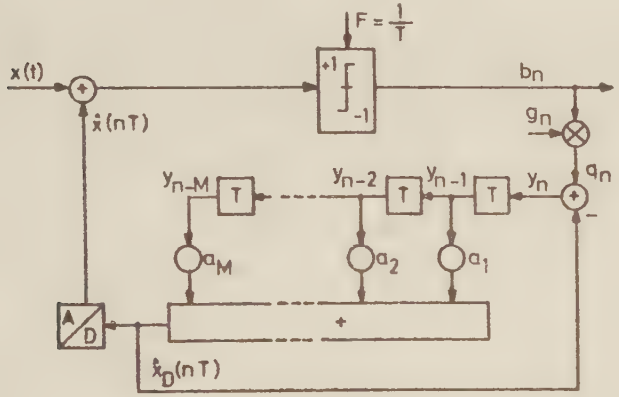


Bild 5.2 Deltamodulator mit Prädiktor zur Grenzzyklusbestimmung

struktionsnetzwerk in der Form eines Transversalfilters nach Gleichung (3.5) mit der Rückkoppelung von $-\hat{x}_n$ ist hier besser geeignet, da jetzt alle benötigten Zustandsgrößen b_n, g_n und y_n für die betrachtete Periodenlänge des Modes direkt zugänglich sind. Die Verwendung des Transversalfilters zur Existenzbestimmung ist zulässig, da die zusätzlichen Nichtlinearitäten, die durch eine eventuell grobe Quantisierung der Reflexionskoeffizienten k_m bzw. der Filterkoeffizienten a_m und durch die endliche Signalwortlänge entstehen, nicht mitbetrachtet werden sollen. Es wird also angenommen, daß Koeffizienten und Signalwerte mit einer ausreichend großen Wortlänge dargestellt werden.

Die Filterkoeffizienten a_m, $m=1,2,\ldots,M$ des Transversalfilters lassen sich eindeutig mit einem rekursiven Algorithmus aus den Reflexionskoeffizienten bestimmen [21]:

$$a_{m,i} = \begin{cases} a_{m-1,i} & , \ i=0 \\ a_{m-1,i} - k_m a_{m-1,m-1} & , \ i=1,2,\ldots,m-1 \\ -k_m & , \ i=m \end{cases} \qquad (5.3)$$

für m = 1, 2, ..., M mit $a_{0,0} = 1$.

Da das Rekonstruktionsnetzwerk stabil sein muß, ist es sinnvoll, die Existenzbedingung nur für solche Koeffizientenkombinationen der a_m zu berechnen, die ein stabiles Rekonstruktionsnetzwerk liefern. Für M > 2 ist diese Stabilitätsbedingung aber einfacher für die Reflexionskoeffizienten anzugeben (Bedingung (3.31)).

Ziel dieser Existenzbetrachtung ist es jetzt also, denjenigen M-dimensionalen Raum zu finden, dessen Koordinaten die Reflexionskoeffizienten k_m, m=1,2,...,M sind mit der Nebenbedingung (3.31), in dem ein bestimmter vorgegebener Mode nicht existieren kann. Für einen bestimmten Mode der Periodenlänge L ergibt sich ein lineares Gleichungssystem der Ordnung L zur Berechnung der Zustandsgrößen

$$y_{n+1} \quad , \qquad 1=0,1,2,\ldots,L-1 \quad . \qquad (5.4)$$

Vorgegeben wird ein zu testender Mode†

$$b_{n+1} \quad , \qquad 1=0,1,2,\ldots,L-1 \quad . \qquad (5.5)$$

Wenn der Deltamodulator eine adaptive Quantisierungsstufeneinstellung (siehe z.B. Tabelle 1) besitzt, bei der die Adaption von g_n von der Folge $\{b_n\}$ abhängt, so muß auch noch die dem jeweiligen Mode entsprechende Folge der $\{g_n\}$ vorgegeben werden. Wenn die Anzahl der Quantisierungsstufen beschränkt ist, muß man je nach Mode dabei zwischen einem Übergangszustand und einem stationären Zustand unterscheiden. Es ist auch für den Übergangszustand noch sinnvoll von einem Grenzzyklus, also einem periodischen Vorgang zu sprechen, da die Folge der

---

†Die b_{n+1} sind jetzt im Gegensatz zur Modendefinition in Abschnitt 5.1 Werte aus {-1,1}.

$\{b_n\}$ auch in diesem Fall periodisch ist. Lediglich die y_n und die g_n sind dann nicht periodisch. Abhängig vom Adaptionsalgorithmus und vom Mode kann man beim eigentlichen stationären Zustand drei Fälle unterscheiden:

1. Die Quantisierungsstufe wird im Übergangszustand im Mittel laufend kleiner (also nicht unbedingt monoton oder streng monoton). Dann arbeitet der Deltamodulator im stationären Zustand mit den kleinsten Quantisierungsstufen.
2. Aufgrund des Adaptionsalgorithmus und des vorgegebenen Modes gibt es keinen Übergangszustand. Die Werte der Zustandsgrößen bei der Anregung eines solchen Modes sind dann rein zufällig, setzen sich aber periodisch fort.
3. Die Quantisierungsstufe wird im Übergangszustand im Mittel laufend größer (also nicht unbedingt monoton oder streng monoton). Dann arbeitet der Deltamodulator im stationären Zustand mit den größten Quantisierungsstufen.

Weiterhin müssen die Koeffizienten k_m bzw. a_m vorgegeben werden. Das lineare Gleichungssystem zur Berechnung der Zustandgrößen y_{n+l} lautet dann für $x(nT)\equiv 0$ (siehe zur Bestimmung auch Bild 5.2):

$$y_{n+l} + \sum_{m=1}^{M} a_m y_{n+l-m} = b_{n+l} h_{n+l} \quad , \quad l=0,1,\ldots,L-1 \qquad (5.6)$$

mit der Periodizitätsbedingung:

$$y_{n+L} = y_n \qquad \forall\, n \; . \qquad (5.7)$$

Wenn die Zustandgrößen y_{n+l} berechnet worden sind, muß folgende Bedingung für die Existenz des zu testenden Modes überprüft werden:

$$b_{n+l} = \operatorname{sgn} \left\{ \sum_{m=1}^{M} a_m y_{n+l-m} \right\} \quad , \quad l=0,1,\ldots,L-1 \qquad (5.8a)$$

oder $\text{sgn}\{y_{n+l}-b_{n+l}g_{n+l}\} = -\text{sgn}\{b_{n+l}g_{n+l}\}$, $l=0,1,\ldots,L-1$. (5.8b)

Wenn diese Bedingung für $l=0,1,\ldots,L-1$ erfüllt ist, dann kann der betreffende Mode existieren.

5.3 Auswahl der zu testenden Grenzzyklusmoden

Mit Hilfe der Existenzbedingung des vorhergehenden Abschnitts ist es jetzt möglich, eine Vielzahl von Moden systematisch zu untersuchen. Wenn man bei dieser Untersuchung aber von vorneherein keine Beschränkungen macht, dann ist die Betrachtung aller Möglichkeiten bei der großen Anzahl der zu variierenden Parameter praktisch unmöglich. Für $x(t)=0$ ist es sicher vernünftig anzunehmen, daß die periodische Folge der Zustandsgrößen y_{n+l} keinen Gleichanteil besitzt. Damit kommen wahrscheinlich nur symmetrische Moden in Frage, bei denen

$$\sum_{l=0}^{L-1} b_{n+l}g_{n+l} = 0 \qquad (5.9)$$

gilt. Es sollen aber auch Moden untersucht werden, bei denen die Symmetriebedingung (5.3) nicht erfüllt ist. Ausgehend von kurzen Periodenlängen sollen also alle symmetrischen Moden für einen Deltamodulator mit einem Rekonstruktionsnetzwerk vom Grad M=2 getestet werden. Für M>2 wird nur eine spezielle Auswahl dieser Moden untersucht. Die symmetrischen Moden bis L = 12 sind in Tabelle 2 wiedergegeben.

Tabelle 2 Symmetrische Grenzzyklusmoden bis Periodenlänge L=12

Eliminiert werden zusätzlich alle Bitmuster, die
1. durch zyklische Vertauschung und/oder Komplementierung der Binärzeichen entstehen,
2. durch periodische Fortsetzung aus Mustern mit kürzerer Periodenlänge entstehen.

L	Bitmuster (Oktal)				
2	1				
4	03				
6	07	13			
8	017	027	033	053	055
10	0037	0057	0067	0073	0127
	0133	0135	0147	0153	0155
	0233	0253	0255	0263	
12	0077	0137	0157	0167	0173
	0257	0267	0273	0275	0317
	0327	0333	0335	0347	0353
	0355	0467	0473	0527	0533
	0535	0547	0553	0555	0563
	0627	0633	0647	0653	0655
	0663	0665	0671	1133	1153
	1253	1255	1263	1265	1315

5.4 Berechnung von Stabilitätsdiagrammen für einen Deltamodulator mit M-stufigem PARCOR-Synthesefilter

Ein Stabilitätsdiagramm für einen bestimmten zu testenden Mode läßt sich am übersichtlichsten im Reflexionskoeffizientenraum darstellen. Bei einem 3-stufigen PARCOR-Filter wird diese Darstellung aber schon kompliziert und die Berechnung sehr aufwendig, da für jede Kombination von Reflexionskoeffizienten und jeden zu testenden Mode ein lineares Gleichungssystem der Ordnung L gelöst werden muß. Zusätzlich muß die Existenzbedingung überprüft werden. Das Stabilitätsprogramm muß dann für eine hinreichend große Anzahl von Reflexionskoeffizientenkombinationen berechnet werden, um die Stabilitätsbereiche genau bestimmen zu können. Bei mehr als drei Reflexionskoeffizienten ist eine anschauliche Darstellung nicht mehr möglich. Aus diesen Gründen wird eine Kombination aus systematischer Berechnung der entsprechenden Stabilitätsdiagramme für zwei Reflexionskoeffizienten und einer parametrischen Darstellung für mehr als zwei Reflexionskoeffizienten durchgeführt.

Die Wertekombinationen für k_m, m=3,4,... werden an den tatsächlich bei der Analyse von Sprachsignalen auftretenden Werten orientiert. Für k_1 und k_2 sind alle Werte zugelassen, die zu einem stabilen Synthesefilter führen.

5.4.1 Stabilitätsdiagramm für 2-stufiges PARCOR-Synthesefilter

Als einfaches aber wichtiges Beispiel für den Test der Existenz eines vorgegebenen Modes soll ein Deltamodulator mit 2-stufigem PARCOR-Filter betrachtet werden. Es ist zu testen, in welchem Bereich der k_1,k_2-Ebene der 4/03-Mode für g_n=const. =1 nicht existieren kann (Stabilitätsbereich). Nach (5.6) ergibt sich folgendes lineare Gleichungssystem zur Bestimmung von y_{n+l}, l=0,1,2,3 und n=0:

$$\begin{aligned} y_o + a_1 y_3 + a_2 y_2 &= -1 \\ y_1 + a_1 y_o + a_2 y_3 &= -1 \\ y_2 + a_1 y_1 + a_2 y_o &= 1 \\ y_3 + a_1 y_2 + a_2 y_1 &= 1 \end{aligned} \qquad (5.10)$$

oder in Matrixschreibweise:

$$\begin{bmatrix} 1 & 0 & a_2 & a_1 \\ a_1 & 1 & 0 & a_2 \\ a_2 & a_1 & 1 & 0 \\ 0 & a_2 & a_1 & 1 \end{bmatrix} \begin{bmatrix} y_o \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} \quad . \tag{5.11}$$

Die Existenzbedingung für den 4/03-Mode lautet:

$$\begin{aligned} y_o + 1 &> 0 \\ y_1 + 1 &> 0 \\ y_2 - 1 &< 0 \\ y_3 - 1 &< 0 \quad . \end{aligned} \tag{5.12}$$

Die Auswertung von 40x40=1600 Reflexionskoeffizientenpaaren ergibt das Stabilitätsdiagramm nach Bild 5.3. Im schraffierten Bereich der k_1, k_2-Ebene ist der 4/03-Mode nicht möglich.

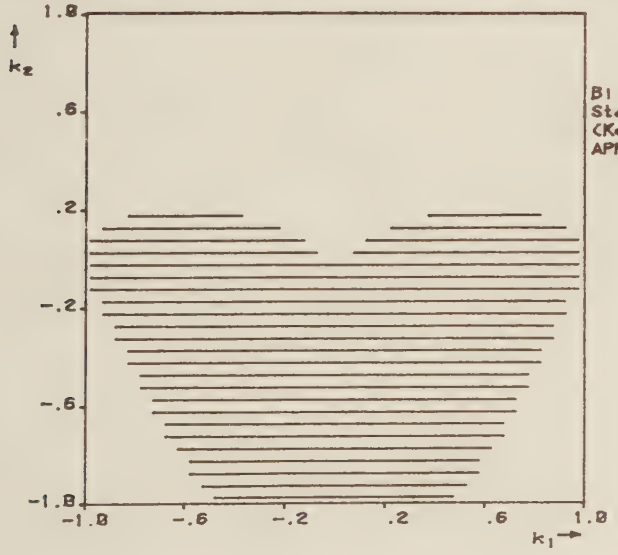


Bild 5.3
Stabilitaetsbereich
(Kein 4/0003-Mode)
APPDM(2)†

† Die Erklärung dieser Kurzbezeichnung findet sich in Anhang A4.

5.4.2 Ermittlung des Reflexionskoeffizientenbereichs bis M=4 bei der PARCOR-Analyse von Sprachsignalen (Abtastrate = 16 kHz und Tiefpaßfilterung bis 3.4 kHz

Bevor weitere Stabilitätsuntersuchungen durchgeführt werden können, ist es sinnvoll, den möglichen Bereich der Reflexionskoeffizienten, so wie er sich bei der Analyse von realen Sprachsignalen ergibt, zu ermitteln. Als Beispiel wird die Sprachprobe eines männlichen Sprechers untersucht. Da die Reflexionskoeffizienten, die im Deltamodulator berechnet werden, mit dem Quantisierungsrauschen behaftet sind, werden diese Koeffizienten dem Betrage nach kleiner sein als die Reflexionskoeffizienten, die durch eine PARCOR-Analyse des gleichen Sprachsignals ermittelt werden. Die Reflexionskoeffizienten des Sprachsignals werden zusammen mit dem Spektrogramm der Sprachprobe dargestellt, damit man entscheiden kann, ob es sich um Pausensegmente oder um Signalsegmente handelt (Bild 5.4).

Für die PARCOR-Analyse wird das Analysefilter nach Bild 3.4 benutzt. Zur Berechnung der Reflexionskoeffizienten wird das Verfahren von Itakura und Saito (Bild 4.7) eingesetzt. Statt $y^+_{n,m-1}$ und $y^-_{n,m-1}$ werden als Eingangssignal für die partielle Korrelation $x^+_{n,m-1}$ und $x^-_{n,m-1}$ benutzt. Der Wert für d in Bild 4.7 beträgt 8, so daß sich eine Zeitkonstante T_i = 16 ms nach Gleichung 4.11 für die Gewichtung ergibt.

Erst durch die Kenntnis des tatsächlich auftretenden Reflexionskoeffizientenbereichs wird es möglich, die Stabilitätsdiagramme entsprechend zu beurteilen. Für das Diagramm von Bild 5.3 bedeutet das, daß der 4/03-Mode auf jeden Fall möglich ist, da gerade im rechten unteren Bereich (k_1 nahe bei 1, k_2 nahe bei -1) diese Reflexionskoeffizienten am häufigsten auftreten. Bei einer Abtastrate von 16 kHz beträgt die Grundfrequenz des 4/03-Modes 4 kHz und liegt damit gerade außerhalb des zu übertragenden Frequenzbereiches. Je nach Tiefpaßfilterung ist dieser Mode dann aber eventuell noch hörbar. Die Existenzmöglichkeit eines Modes bedeutet aber nicht, daß dieser Mode auch auftreten muß. Die Bedingungen für das tatsächliche Auftreten eines Modes werden weiter unten diskutiert werden.

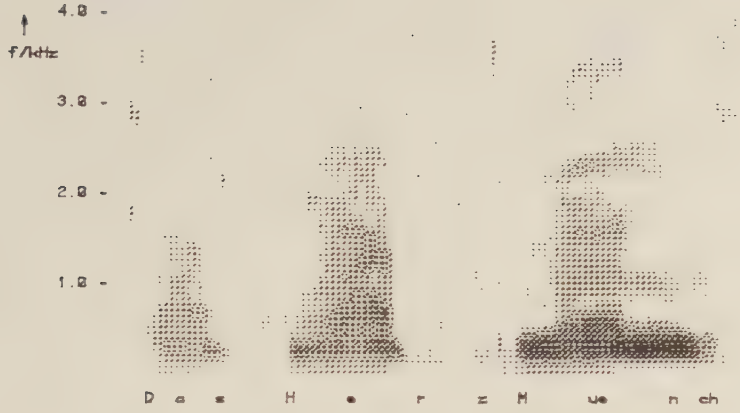


Bild 5.4a Spektrogramm/Datei: M116M /Teil 1
Dynamik= 50 dB

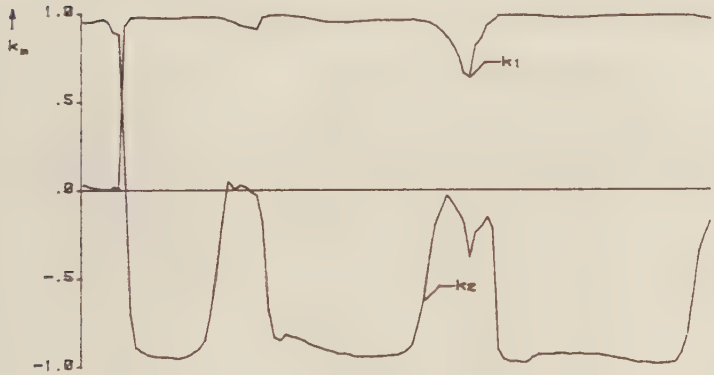


Bild 5.4b Reflexionskoeffizientenverlauf

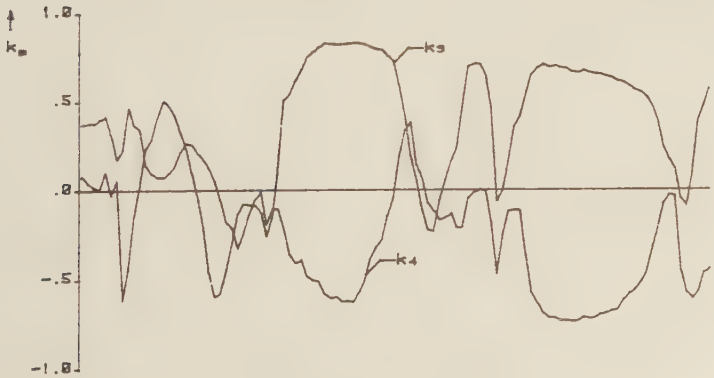


Bild 5.4c Reflexionskoeffizientenverlauf

5.4.3 Untersuchung einiger ausgewählter unsymmetrischer Moden

Da es nicht von vorneherein ausgeschlossen ist, daß auch unsymmetrische Moden bei $x(t)\equiv 0$ existieren können, müssen auch diese zumindest teilweise betrachtet werden. Wenn unsymmetrische Moden existieren können, dann hängt das sicher vom Grad der Unsymmetrie ab. Es wurden deshalb Moden mit

$$\left| \sum_{l=0}^{L-1} b_{n+1} \right| = 1$$

und

$$\left| \sum_{l=0}^{L-1} b_{n+1} \right| = 2$$

für $g_{n+1}=1$ untersucht. Für M=2 können z.B. der 3/3-Mode und

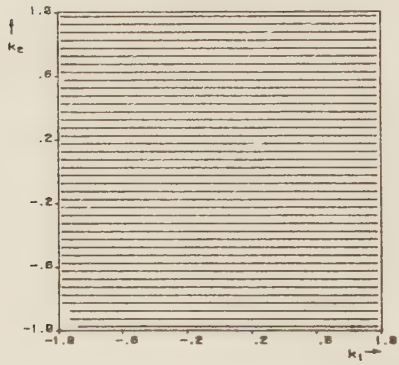


Bild 5.5a
Stabilitätsbereich
(Kein 3/0003-Mode)
APPDM(2)

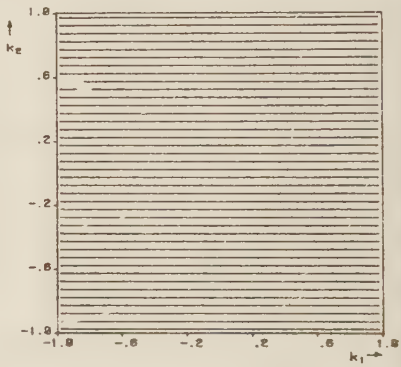


Bild 5.5b
Stabilitätsbereich
(Kein 5/0007-Mode)
APPDM(2)

der 5/07-Mode existieren (Bild 5.5). Die Existenzbereiche für diese Moden liegen allerdings so, daß sie normalerweise von k_1 und k_2 (Bild 5.4b) nicht erreicht werden.

Moden mit einem noch größeren Grad an Unsymmetrie sind z.B. der 4/07- und der 6/17-Mode. Für M=2 stellte es sich heraus, daß der 4/07-Mode im untersuchten Bereich nicht existieren kann. Für den 6/17-Mode (Bild 5.6a) gelten die gleichen Aussagen wie für den 3/3- und den 5/07-Mode.

Für den 6/17-Mode wurden auch noch Untersuchungen für

M=4 durchgeführt, wobei die Wertepaare für k_3 und k_4 aufgrund von Bild 5.4c ausgewählt wurden. Für die Wertepaare

$$k_3 = 0.2,\ k_4 = 0.2$$
$$k_3 = 0.2,\ k_4 = -0.2$$
$$k_3 = -0.2,\ k_4 = 0.2$$
$$k_3 = -0.2,\ k_4 = -0.2$$

ergaben sich gegenüber Bild 5.6a keine Veränderungen. Für k_3=0.4 und k_4=-0.4, eine Wertekombination, die nach Bild 5.4c sehr wahrscheinlich ist, ist der 6/17-Mode im untersuchten Bereich nicht möglich. Für die Kombination k_3=0.4, k_4=0.4, die kaum vorkommen kann (Bild 5.6b), ergab sich lediglich eine geringfügige Änderung gegenüber einem Deltamodulator mit 2-stufigem PARCOR-Filter.

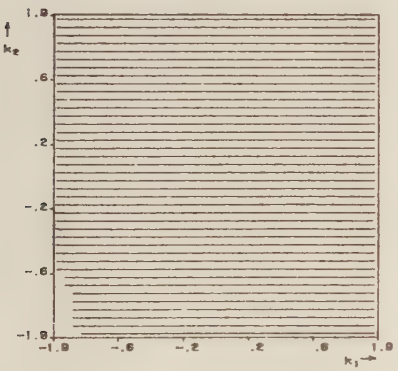


Bild 5.6a
Stabilitätsbereich
(Kein 6/8817-Mode)
APPDN(2)

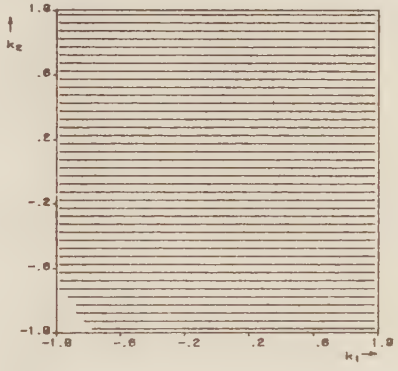


Bild 5.6b
Stabilitätsbereich
(Kein 6/8817-Mode) k_3= -.4
APPDN(4) k_4= .4

5.4.4 Systematische Untersuchung aller symmetrischen Moden bis zur Periodenlänge L=12 bei 2-stufigem PARCOR-Synthesefilter und fester Quantisierungsstufe (g_n=1)

Um sich einen Überblick über die Größe und Verteilung der Existenzbereiche der möglichen Grenzzyklen zu verschaffen, ist es zunächst sinnvoll, alle symmetrischen Moden bis zu einer maximalen Periodenlänge zu untersuchen, über die hinaus wahrscheinlich kaum noch Grenzzyklen existieren können. Bei der Betrachtung dieser Existenzbereiche sollte auch der tatsächlich mögliche Bereich der Reflexionskoeffizienten berücksichtigt werden (siehe auch Abschnitt 5.4.2).

Nach Tabelle 2 ergeben sich bis zur Periodenlänge L=12 63 verschiedene zu testende symmetrische Moden. Von diesen Moden sind im untersuchten Bereich

$$-1 < k_1 < 1 \quad , \quad -1 < k_2 < 1$$

nur die in den Bildern 5.7, 5.8 und 5.9 (4/03-Mode in Bild 5.3) dargestellten Moden existenzfähig. Wenn man darüber hinaus den tatsächlich möglichen Bereich für k_1 und k_2 betrachtet, so sind nur die folgenden Moden existenzfähig:

2/1 -Mode ,
4/03 -Mode ,
6/07 -Mode ,
6/13 -Mode ,
8/053 -Mode ,
10/0253-Mode ,
10/0263-Mode .

Der 2/1-Mode, als Grundmode des Deltamodulators, darf natürlich immer existieren und stört auch nicht. Der 4/03-Mode wurde schon in Abschnitt 5.4.2 diskutiert.

Berücksichtigt man weiterhin, daß im Deltamodulator aufgrund des Quantisierungsgeräusches erfahrungsgemäß gilt†:

$$k_1 < 0.995$$
$$k_2 > -0.9 \quad ,$$

†F = 15 kHz.

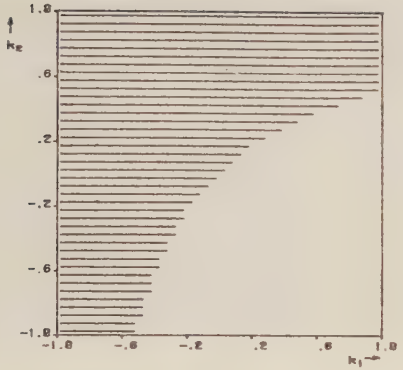


Bild 5.7a
Stabilitätsbereich
(Kein 2/0001-Mode)
APPDM(2)

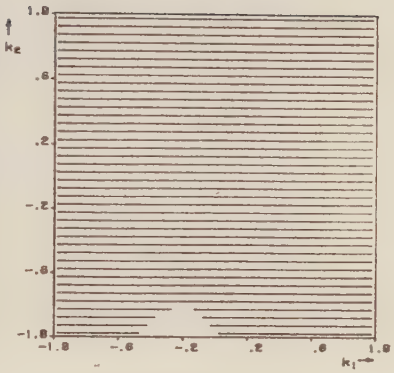


Bild 5.7d
Stabilitätsbereich
(Kein 6/0017-Mode)
APPDM(2)

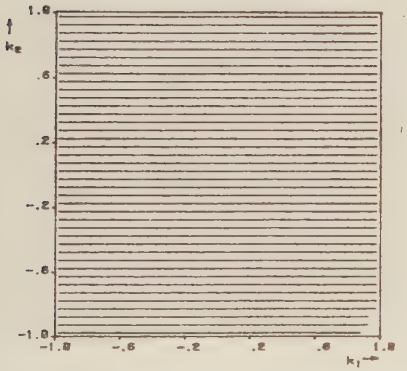


Bild 5.7b
Stabilitätsbereich
(Kein 6/0007-Mode)
APPDM(2)

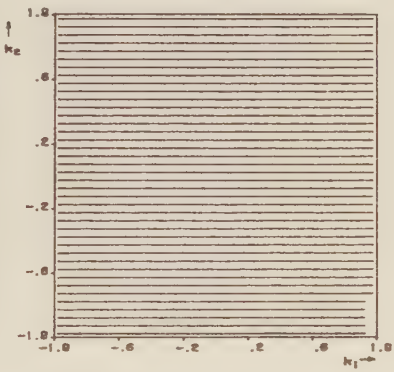


Bild 5.7e
Stabilitätsbereich
(Kein 6/0053-Mode)
APPDM(2)

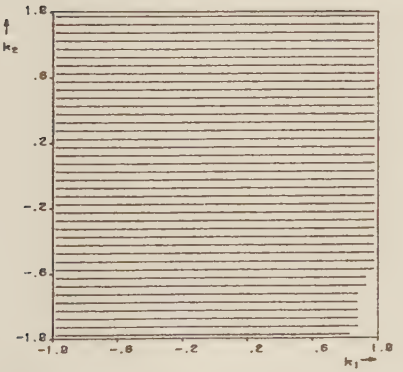


Bild 5.7c
Stabilitätsbereich
(Kein 6/0013-Mode)
APPDM(2)

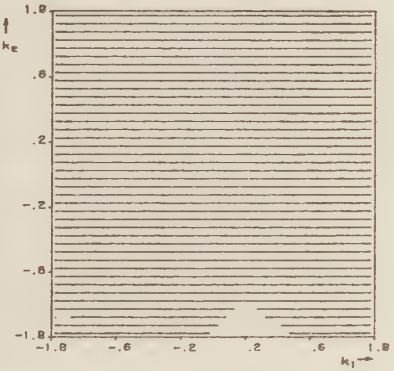


Bild 5.7f
Stabilitätsbereich
(Kein 6/0055-Mode)
APPDM(2)

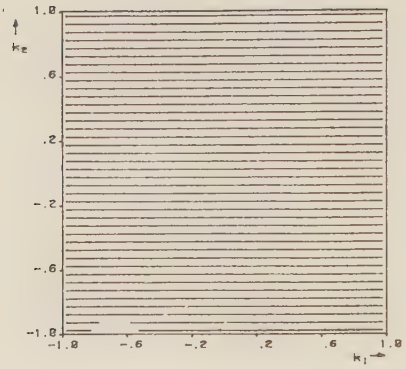


Bild 5.6a
Stabilitätsbereich
(Kein 10/0037-Mode)
APPDM(2)

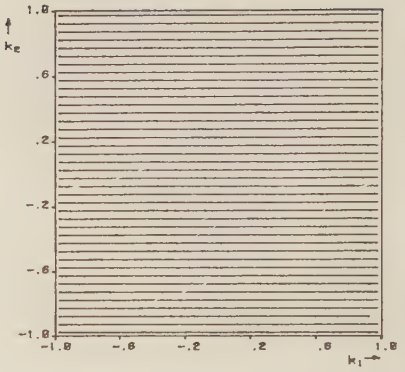


Bild 5.6d
Stabilitätsbereich
(Kein 10/0253-Mode)
APPDM(2)

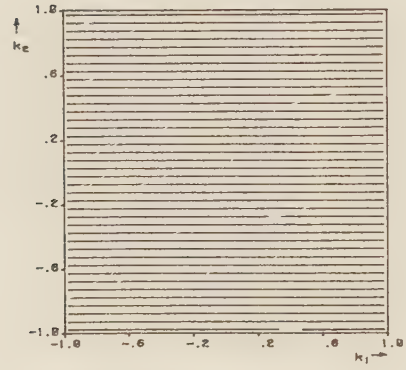


Bild 5.6b
Stabilitätsbereich
(Kein 10/0135-Mode)
APPDM(2)

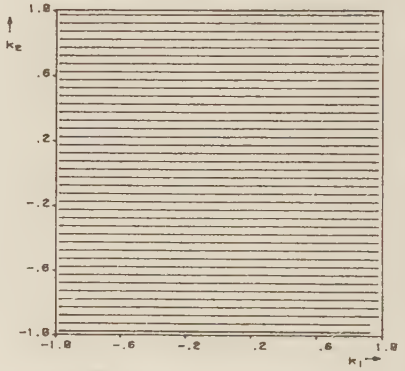


Bild 5.6e
Stabilitätsbereich
(Kein 10/0283-Mode)
APPDM(2)

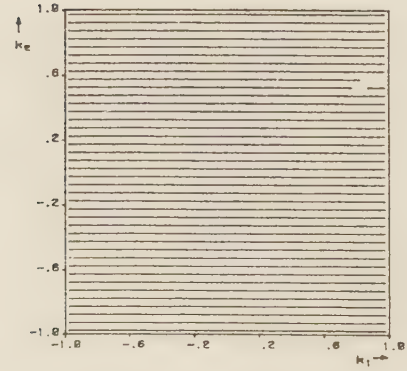


Bild 5.6c
Stabilitätsbereich
(Kein 10/0233-Mode)
APPDM(2)

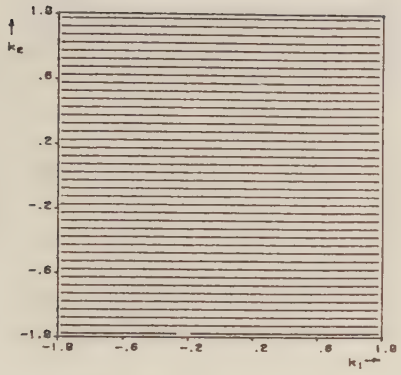


Bild 5.8a
Stabilitätsbereich
(Kein 12/0257-Mode)
APPDN(2)

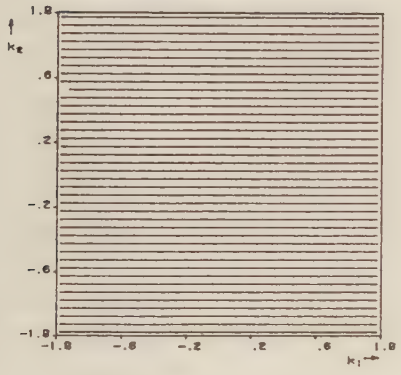


Bild 5.8b
Stabilitätsbereich
(Kein 12/0665-Mode)
APPDN(2)

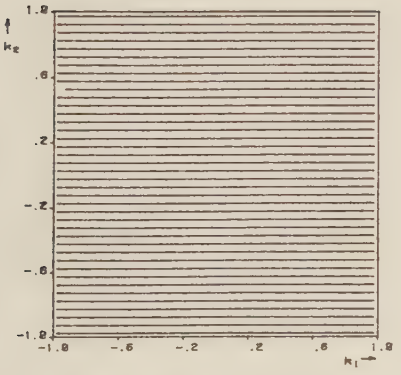


Bild 5.8c
Stabilitätsbereich
(Kein 12/1265-Mode)
APPDN(2)

dann sind die verbleibenden Existenzbereiche für den

6/07 -Mode ,
8/053 -Mode ,
10/0253-Mode
und 10/0263-Mode

sehr klein (der Abstand zwischen zwei Linien in den Stabilitätsdiagrammen beträgt 0.05). Damit ist nur noch der 6/13-Mode von einer gewissen Bedeutung. Zusammenfassend kann man sicher sagen, daß in einem Deltamodulator mit 2-stufigem PARCOR-Filter das Entstehen von störenden Grenzzyklen nicht sehr wahrscheinlich ist.

5.4.5 Untersuchung der Änderung der Stabilitätsbereiche für einige ausgewählte symmetrische Moden bei 4-stufigem PARCOR-Synthesefilter mit fester Quantisierungsstufe (g_n=1)

Eine systematische Untersuchung aller symmetrischen Grenzzyklusmoden für höhere Rekonstruktionsfiltergrade (M>2) ist praktisch nicht durchführbar. Es wird deshalb als Beispiel dieVeränderung der Stabilitätsbereiche für M=4 bei drei ausgewählten Moden betrachtet. Wenn keine sprunghaften Veränderungen dieser Bereiche bei stetiger Veränderung der Parameter auftreten, so kann man daraus Schlüsse auch für die Veränderung der Stabilitätsbereiche anderer Moden ziehen. Die drei Moden, die untersucht werden sollen, sind:

6/07 - Mode ,
8/053 - Mode
und 8/055 - Mode .

Die ausgewählten Parameterkombinationen für k_3 und k_4 sind in Tabelle 3 dargestellt. Das ist natürlich nur eine begrenzte Anzahl von Wertepaaren. Werte von $|k_3|$ und $|k_4| > 0.5$ wurden nicht betrachtet, da bekannt ist, daß der Betrag der Reflexionskoeffizienten k_i, i=3,4,... für Abtastraten von ca. 10 kHz kleiner als 0.7 ist $|21|$. Durch das Quantisierungsgeräusch werden diese Werte noch kleiner, so daß diese Einschränkung auch für Abtastraten bis 16 kHz realistisch erscheint.

Tabelle 3 Wertepaare für k_3 und k_4

k_3	k_4
0.1	0.1
0.1	-0.1
-0.1	0.1
-0.1	-0.1
0.2	0.2
0.2	-0.2
-0.2	0.2
-0.2	-0.2
0.3	-0.3
0.4	-0.4
0.5	-0.5
0.6	-0.6

Nach Bild 5.4c sind für kleine Werte von k_3 und k_4 alle Kombinationen möglich. Bei größeren Werten ($k_3 \geq 0.3$) gilt meistens:

$$k_3 \approx -k_4 \quad .$$

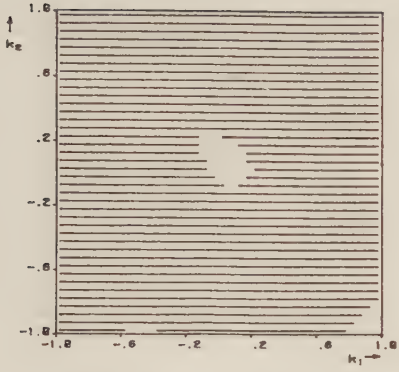


Bild 5.10a
Stabilitätsbereich
(Kein 6/8887-Mode) k_3= .1
APPDN(4) k_4= .1

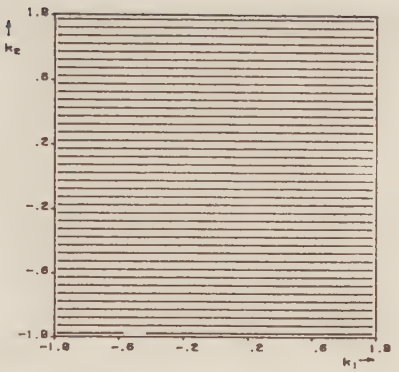


Bild 5.10d
Stabilitätsbereich
(Kein 6/8887-Mode) k_3= -.1
APPDN(4) k_4= -.1

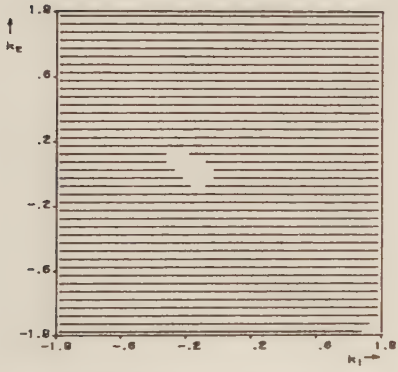


Bild 5.10b
Stabilitätsbereich
(Kein 6/8887-Mode) k_3= .1
APPDN(4) k_4= .1

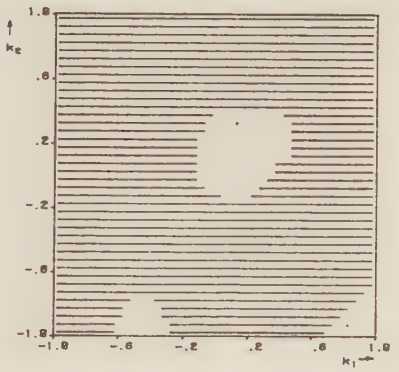


Bild 5.10e
Stabilitätsbereich
(Kein 6/8887-Mode) k_3= .2
APPDN(4) k_4= .2

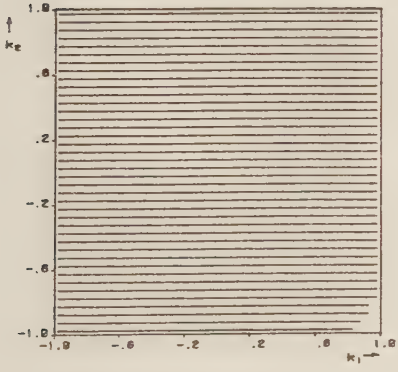


Bild 5.10c
Stabilitätsbereich
(Kein 6/8887-Mode) k_3= -.1
APPDN(4) k_4= .1

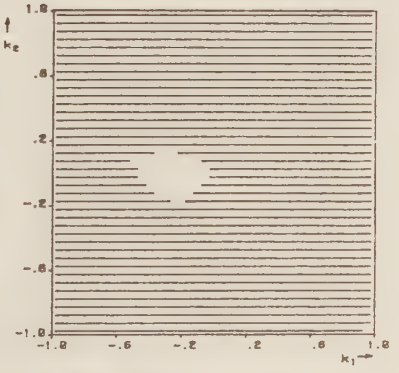


Bild 5.10f
Stabilitätsbereich
(Kein 6/8887-Mode) k_3= .2
APPDN(4) k_4= -.2

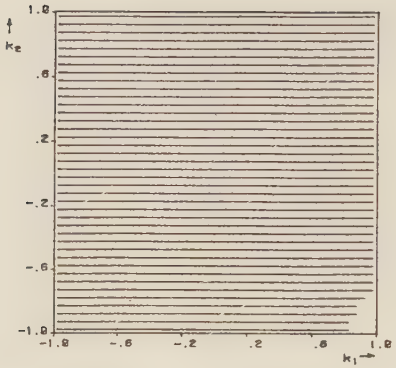


Bild 5.18g
Stabilitätsbereich
(Kein 6/8887-Mode) k_3= -.2
APPDM(4) k_4= .2

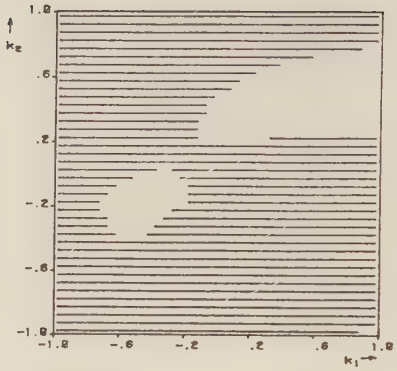


Bild 5.18j
Stabilitätsbereich
(Kein 6/8887-Mode) k_3= .4
APPDM(4) k_4= -.4

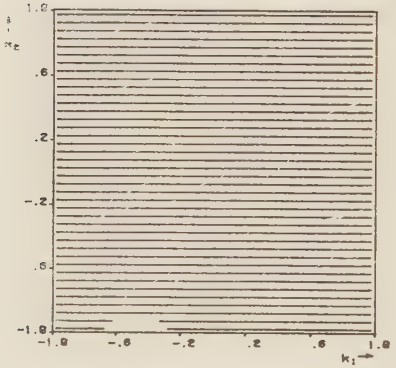


Bild 5.18h
Stabilitätsbereich
(Kein 6/8887-Mode) k_3= -.2
APPDM(4) k_4= -.2

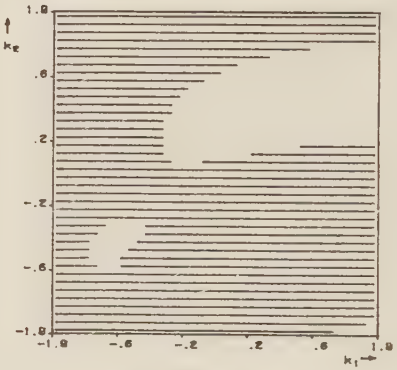


Bild 5.18k
Stabilitätsbereich
(Kein 6/8887-Mode) k_3= .5
APPDM(4) k_4= -.5

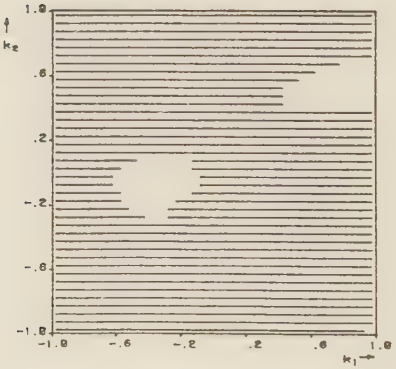


Bild 5.18i
Stabilitätsbereich
(Kein 6/8887-Mode) k_3= .3
APPDM(4) k_4= -.3

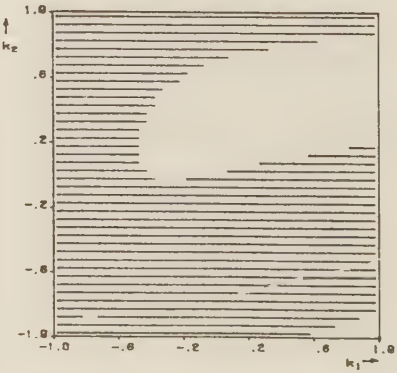


Bild 5.18l
Stabilitätsbereich
(Kein 6/8887-Mode) k_3= .6
APPDM(4) k_4= -.6

Die Ergebnisse dieser Untersuchung zeigt das Bild 5.10 und im Anhang die Bilder A5.1 und A5.2. Diese Diagramme zeigen, daß sich gerade im kritischen Bereich (in der rechten unteren Ecke) der k_1, k_2-Ebene keine großen Veränderungen gegenüber einem Deltamodulator mit 2-stufigem PARCOR-Synthesefilter ergeben. Bei bestimmten Wertepaaren von k_3 und k_4 werden die entsprechenden Moden in diesem Bereich sogar nicht existenzfähig. Größere Veränderungen der Stabilitätsbereiche ergeben sich nur bei großen Beträgen von k_3 und k_4. Beim 6/07-Mode kommen bei großen Beträgen von k_3 und k_4 nur dort größere Instabilitätsbereiche hinzu, wo k_1 und k_2 normalerweise nicht auftreten.

5.4.6 Änderung der Stabilitätsbereiche einiger ausgewählter Moden bei 2- und 4-stufigem PARCOR-Synthesefilter und adaptiver Quantisierungsstufeneinstellung

Bei einem Deltamodulator mit Augenblicksadaption der Quantisierungsstufe $g_n=f(b_n,b_{n-1},...)$ ist es möglich, für einen vorgegebenen Mode, den Wert der Quantisierungsstufe für einen bestimmten Adaptionsalgorithmus eindeutig zu jedem Abtastzeitpunkt anzugeben. Abhängig vom Verhältnis k aufeinanderfolgender Quantisierungsstufen wird sich dann der Stabilitätsbereich für den betreffenden Mode mehr oder weniger verändern. Dabei ist es unter Umständen möglich, daß sich durch die Einführung dieser Adaption die Existenzbereiche für bestimmte Moden sogar verkleinern.

Eine gewisse zusätzliche Schwierigkeit bei dieser Untersuchung besteht darin, daß man zwischen einem Übergangsbereich, in dem die Adaption voll wirksam ist, und einem stationären Bereich unterscheiden muß (siehe auch Abschnitt 5.2). Besonders kritisch sind dabei solche Moden, bei denen auf grund des Adaptionsalgorithmus die größten Quantisierungsstufen eingeschaltet werden oder bei denen die Quantisierungsstufen zufällig groß sind und sich nicht ändern. Solche Moden sollten nach Möglichkeit auf keinen Fall existieren dürfen.

Es soll nur der stationäre Fall genauer untersucht werden. Als Beispiel werden dazu der

4/o3 - Mode ,
6/07 - Mode
und der 6/13 - Mode

betrachtet. Als besonders günstig bei der Codierung von Sprachsignalen bei niedrigen Abtastraten haben sich die Adaptionsalgorithmen mach Tabelle 1 erwiesen, bei denen die Adaption von den letzten beiden oder den letzten drei übertragenen Binärwerten b_n abhängt. Der Fall, bei dem die Adaption von drei vergangenen Binärwerten abhängt, wurde aber aus folgenden Gründen nicht genauer betrachtet:

1. Wenn im zu testenden Mode nicht mehr als zwei gleiche Binärwerte aufeinanderfolgen, schaltet der Adaptionsalgorithmus im stationären Zustand mit Ausnahme des 4/03-Modes die kleinsten Quantisierungsstufen ein. Es ergibt sich sogar teilweise eine konstante Quantisierungsstufe, so daß sich die Existenzbedingungen dieser Moden gegenüber dem nichtadaptiven Fall nicht ändern.
2. Wenn drei oder mehr aufeinanderfolgende Binärwerte bei den symmetrischen Moden bis L=12 gleich sind, ändert sich das lineare Gleichungssystem (5.6), wie sich bei eingehender Untersuchung zeigen läßt, kaum. Damit werden sich die Stabilitätsdiagramme für diese Moden ebenfalls kaum ändern.

Die gleichen Argumente gelten sinngemäß noch stärker bei einer Adaption, die von mehr als drei vergangenen Binärwerten abhängig ist. Zu beachtende Änderungen in den Stabilitätsdiagrammen treten also nur bei einer Adaption der Quantisierungsstufe durch zwei vergangene Binärwerte auf (wenn der Adaptionsfaktor k nicht zu groß wird). Untersucht werden dabei nur die symmetrischen Moden bis L=6, da darüber hinaus entweder nur die kleinsten Quantisierungsstufen eingeschaltet werden oder aber die linearen Gleichungssysteme und die Existenzbedingungen bei Moden mit vielen aufeinanderfolgenden gleichen Binärzeichen kaum geändert werden. Die veränderten Gleichungssysteme und Existenzbedingungen für die oben angegebenen Moden zeigt Tabelle 4.

Die Ergebnisse dieser Stabilitätsuntersuchung für M=2

Tabelle 4 Lineare Gleichungssysteme und Existenzbedingungen für Grenzzyklen bei Augenblicksadaption nach Tabelle 1 (stationärer Fall, n=0). Kurzbezeichnung für Deltamodulator: APPDM(M)/(2/2/k)
Es gilt für y_{l-m}, $l=0,1,\ldots,L-1$:

$$l-m < 0 \rightarrow y_{l-m} = y_{l-m+L}$$

4/03 - Mode ($l=0,1,2,3$; $L=4$): Die Quantisierungsstufen haben zufällige Werte.

$$y_o + \sum_{m=1}^{M} a_m y_{o-m} = -g_{min} k^i \quad , \quad y_o + g_{min} k^i > 0$$

$$y_1 + \sum_{m=1}^{M} a_m y_{1-m} = -g_{min} k^{i+1} \quad , \quad y_1 + g_{min} k^{i+1} > 0$$

$$y_2 + \sum_{m=1}^{M} a_m y_{2-m} = g_{min} k^i \quad , \quad y_2 - g_{min} k^i < 0$$

$$y_3 + \sum_{m=1}^{M} a_m y_{3-m} = g_{min} k^{i+1} \quad , \quad y_3 - g_{min} k^{i+1} < 0$$

6/07 - Mode ($l=0,1,2,3,4,5$; $L=6$): Die Quantisierungsstufen erreichen die maximalen Werte.

$$y_o + \sum_{m=1}^{M} a_m y_{o-m} = -g_{min} k^{i_{max}-1} \quad , \quad y_o + g_{min} k^{i_{max}-1} > 0$$

$$y_1 + \sum_{m=1}^{M} a_m y_{1-m} = -g_{min} k^{i_{max}} \quad , \quad y_1 + g_{min} k^{i_{max}} > 0$$

$$y_2 + \sum_{m=1}^{M} a_m y_{2-m} = -g_{min} k^{i_{max}} \quad , \quad y_2 + g_{min} k^{i_{max}} > 0$$

$$y_3 + \sum_{m=1}^{M} a_m y_{3-m} = g_{min} k^{i_{max}-1} \quad , \quad y_3 - g_{min} k^{i_{max}-1} < 0$$

$$y_4 + \sum_{m=1}^{M} a_m y_{4-m} = g_{min} k^{i_{max}} \quad , \quad y_4 - g_{min} k^{i_{max}} < 0$$

$$y_5 + \sum_{m=1}^{M} a_m y_{5-m} = g_{min} k^{i_{max}} \quad , \quad y_5 - g_{min} k^{i_{max}} < 0$$

Tabelle 4 (Fortsetzung)

6/13 - Mode (l=0,1,2,3,4,5; L=6): Die Quantisierungsstufen erreichen die minimalen Werte.
$y_o + \sum_{m=1}^{M} a_m y_{o-m} = -g_{min}$, $y_o + g_{min} > 0$
$y_1 + \sum_{m=1}^{M} a_m y_{1-m} = -g_{min}\,k$, $y_1 + g_{min}\,k > 0$
$y_2 + \sum_{m=1}^{M} a_m y_{2-m} = g_{min}$, $y_2 - g_{min} < 0$
$y_3 + \sum_{m=1}^{M} a_m y_{3-m} = -g_{min}$, $y_3 + g_{min} > 0$
$y_4 + \sum_{m=1}^{M} a_m y_{4-m} = g_{min}$, $y_4 - g_{min} < 0$
$y_5 + \sum_{m=1}^{M} a_m y_{5-m} = g_{min}\,k$, $y_5 - g_{min}\,k < 0$

zeigen die Bilder 5.11 und 5.12 sowie im Anhang Bild A6.1. Für die betrachteten Adaptionsfaktoren k ändern sich die Stabilitätsdiagramme in den wesentlichen Bereichen kaum. Beim 6/13-Mode mit k=1.3 (Bild A6.1c) wird der Existenzbereich sogar kleiner. Für M=4, k=1.2 und einige ausgewählte Koeffizientenpaare k_3, k_4 sind die Stabilitätsdiagramme in Bild 5.13 und im Anhang in den Bildern A7.1 und A7.2 dargestellt.

Besonders kritisch ist der 6/07-Mode, da bei diesem Mode auf jeden Fall die größten Quantisierungsstufen eingeschaltet werden und für eine Abtastrate von 16 kHz die Grundschwingung dieses Modes von $2\frac{2}{3}$ kHz im Übertragungsbereich liegt. Für k_3= 0.2 und k_4=0.2 wird der Existenzbereich im kritischen Bereich der k_1, k_2-Ebene stark vergrößert. Für alle anderen betrachteten Koeffizientenpaare k_3, k_4 bleibt der Existenzbereich im wesentlichen erhalten. Die Kombination k_3=0.5 und k_4=-0.5 wird in einem Deltamodulator erfahrungsgemäß kaum auftreten. Zusätzlich ist allgemein auch noch zu sagen, daß nicht beliebige Koeffizientenpaare k_1, k_2 im kritischen Bereich auftre-

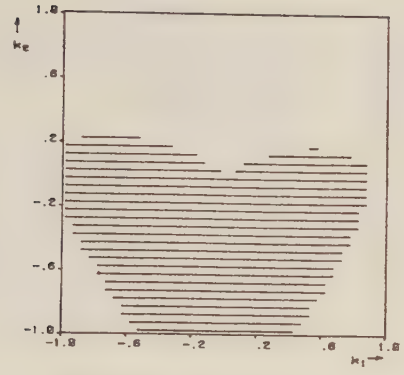


Bild 5.11a
Stabilitaetsbereich
(Koln 4/0003-Mode)
APPDM(2)/(2/2/1.18)

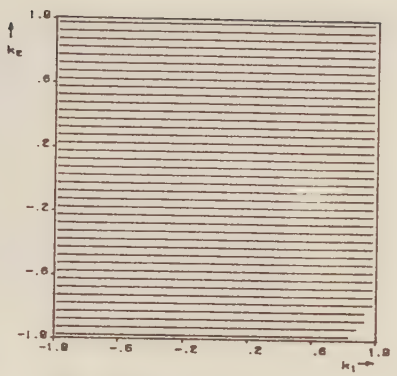


Bild 5.12a
Stabilitaetsbereich
(Koln 6/0007-Mode)
APPDM(2)/(2/2/1.18)

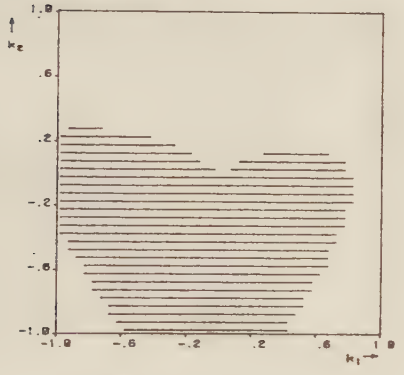


Bild 5.11b
Stabilitaetsbereich
(Koln 4/0003-Mode)
APPDM(2)/(2/2/1.28)

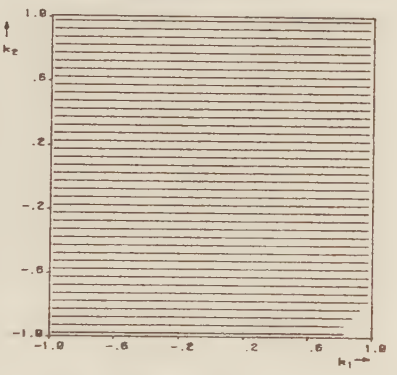


Bild 5.12b
Stabilitaetsbereich
(Koln 6/0007-Mode)
APPDM(2)/(2/2/1.28)

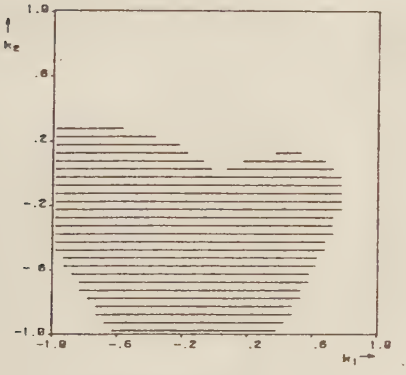


Bild 5.11c
Stabilitaetsbereich
(Koln 4/0003-Mode)
APPDM(2)/(2/2/1.38)

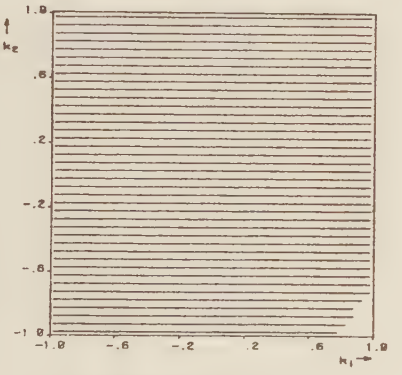


Bild 5.12c
Stabilitaetsbereich
(Koln 6/0007-Mode)
APPDM(2)/(2/2/1.38)

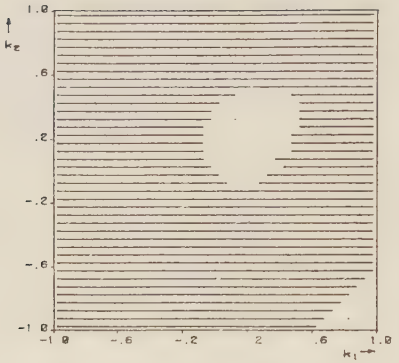


Bild 5.13a
Stabilitaetsbereich
(Koln 6/0007-Mode) k_3= .2
APPDM(4)/(2/2/1.28) k_4= .2

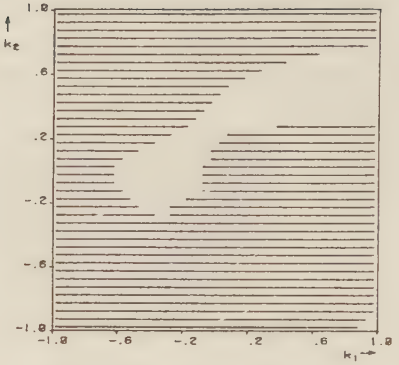


Bild 5.13d
Stabilitaetsbereich
(Koln 6/0007-Mode) k_3= .3
APPDM(4)/(2/2/1.28) k_4= -.3

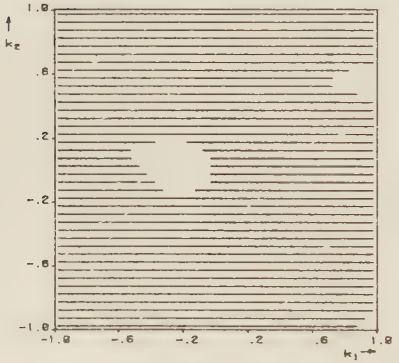


Bild 5.13b
Stabilitaetsbereich
(Koln 6/0007-Mode) k_3= .2
APPDM(4)/(2/2/1.28) k_4= -.2

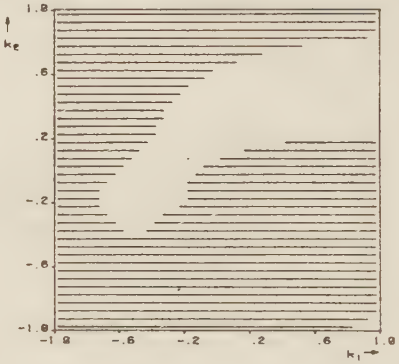


Bild 5.13e
Stabilitaetsbereich
(Koln 6/0007-Mode) k_3= .4
APPDM(4)/(2/2/1.28) k_4= -.4

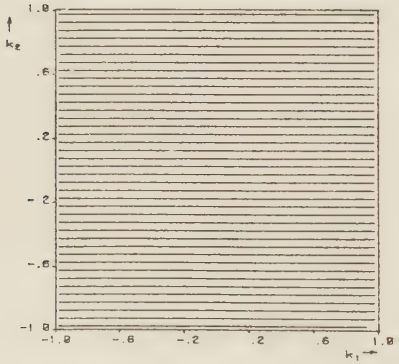


Bild 5.13c
Stabilitaetsbereich
(Koln 6/0007-Mode) k_3= -.2
APPDM(4)/(2/2/1.28) k_4= -.2

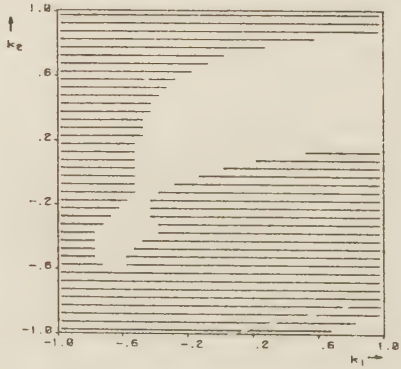


Bild 5.13f
Stabilitaetsbereich
(Koln 6/0007-Mode) k_3= .5
APPDM(4)/(2/2/1.28) k_4= -.5

ten. Der Koeffizient k_2 ist dem Betrage nach für große Werte von k_1 immer kleiner als k_1. Die Veränderungen der Stabilitätsdiagramme in Bild 5.13d bis 5.13f gegenüber dem nichtadaptiven Fall sind ebenfalls nicht kritisch.

5.5 Existenzbedingung für beliebig vorgebbare Moden bei einem Deltamodulator mit M-stufigem PARCOR-Synthesefilter und adaptiver Quantisierungsstufeneinstellung bei x(t)=const.≠0

Die Betrachtung der Existenzbedingung für beliebig vorgebbare Grenzzyklusmoden bei konstantem von Null verschiedenem Eingangssignal ist dann interessant, wenn Gleichspannungsfehler x_g beim Eingangssignal, im Komparator (1-bit-Quantisierer) oder im D/A-Wandler auftreten. Es wird dann, wenn nur dieser Gleichspannungsfehler im System wirksam ist, eine periodische Folge von Rekonstruktionswerten y_n auftreten, die um den Gleichspannungsfehler pendeln. Im allgemeinen ist nicht zu erwarten, daß der Mittelwert $\bar{y}$ von y_n gleich dem Gleichspannungsfehler x_g ist. Eine mögliche periodische Folge bei nichtadaptivem Deltamodulator zeigt Bild 5.14. Es

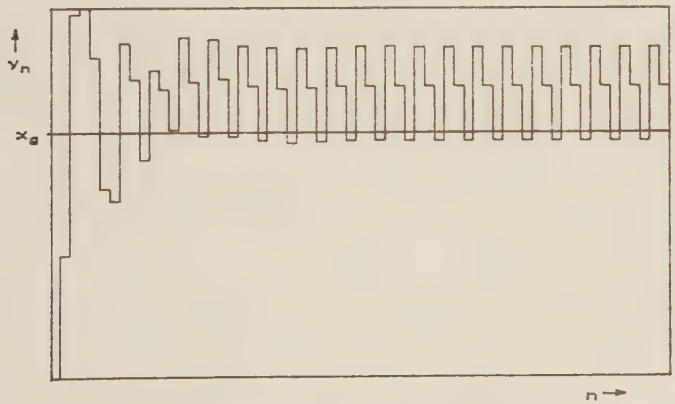


Bild 5.14 Sprungantwort APPDMC 4) fuer
Gleichspannungsfehler $x_g/g_n=2.0$
$k_1=.90$, $k_2=-.70$, $k_3=.40$, $k_4=-.40$

stellt sich also ein Regelfehler e zwischen $\bar{y}$ und x_g ein. Dieser Regelfehler ist zufällig und hängt von den Anfangsbedingungen im Deltamodulator ab. Für jeden Mode gibt es daher

bei konstantem x_g einen bestimmten Bereich für den Regelfehler und den Mittelwert $\bar{y}$, so daß gilt:

$$x_g = \bar{y} + e \quad . \tag{5.13}$$

Will man also die Existenz eines bestimmten Modes für einen bestimmten Gleichspannungsfehler x_g berechnen, dann muß man $\bar{y}$ und e so vorgeben, daß (5.13) gilt. Das lineare Gleichungssystem (siehe auch Gleichungssystem 5.6) zur Bestimmung der

$$\hat{y}_n = y_n - \bar{y} \quad ,$$

wobei $\hat{y}_n$ der gleichspannungsfreie Anteil ist, lautet dann:

$$\hat{y}_{n+1} + \sum_{m=1}^{M} a_m \hat{y}_{n+1-m} = b_{n+1} g_{n+1} - \bar{y} \sum_{m=0}^{M} a_m \quad , \; l=0,1,..,L-1 \tag{5.14}$$

mit $a_o = 1$.

Die Existenzbedingung hat dann folgende Gestalt:

$$\mathrm{sgn}(\hat{y}_{n+1} - b_{n+1} g_{n+1} - e) = -\mathrm{sgn}(b_{n+1} g_{n+1}), \; l=0,1,..,L-1. \tag{5.15}$$

Die Vielfalt der Möglichkeiten der Existenz von Grenzzyklen ist hier noch größer als im Fall $x(t) \equiv 0$, da hier $\bar{y}$ und e vorgegeben werden können. In Abhängigkeit dieser beiden vorzugebenden Größen ändern sich die Stabilitätsbereiche der einzelnen vorzugebenden Moden.

Es sollen jetzt hier keine Stabilitätsdiagramme solcher Moden dargestellt werden. Es wurden allerdings einige Fälle untersucht. Dabei ist allgemein festzustellen, daß sich in Abhängigkeit vom wachsenden Verhältnis $|x_g/g_n|$ die Existenzmöglichkeit von Moden mit beliebig großer Periodenlänge stark vergrößert. Es sind jetzt nicht mehr nur symmetrische Moden zu betrachten, sondern auch beliebig unsymmetrische Moden.

Das bedeutet aber, daß die Grundschwingung dieser Moden bei niedrigen Abtastraten auf jeden Fall im Übertragungsbereich liegt. Bei entsprechenden Adaptionsalgorithmen der Quantisierungsstufe g_n sind auch die y_n entsprechend groß, so daß solche Moden sehr stark stören können. Es ist also

bei der Realisierung eines solchen Deltamodulators darauf zu achten, daß die Gleichspannungssignale (niedrige Frequenzen!) im System so gering wie möglich bleiben.

5.6 Anregung von Grenzzyklen

In der Literatur |43-45| wird häufig darüber berichtet, daß bei Quellcodierungsverfahren, die nach dem Prinzip der adaptiven Prädiktion (APC) arbeiten, Instabilitäten bei grober Quantisierung des Differenzsignals beobachtet wurden. Da bei diesen Verfahren eine niedrige Übertragungsbitrate angestrebt wurde, wurden auch dort nur 1- oder 2-bit Quantisierer eingesetzt. Der adaptive Prädiktor wurde aber aus dem Originalsignal berechnet,und die Prädiktorkoeffizienten wurden getrennt zum Empfänger übertragen. In |24| und |43| wurden auf diese Weise keine größeren Prädiktorgrade als 2 bis 4 erreicht bzw. es wurde darauf hingewiesen, daß diese Verfahren bei grober Quantisierung nicht mehr anzuwenden wären.

In |44| wurde dagegen die Verstärkung des aus dem Originalsignal berechneten Prädiktors künstlich erniedrigt, um die Stabilität des Sytems zu gewährleisten.

Da solche prädiktiven Codiersysteme aber alle rückgekoppelt sind, wird auch der - nach der RDT - unvermeidliche Quantisierungsfehler rückgekoppelt und entsprechend verstärkt. Wenn in einem solchen System die Existenzbedingung für Grenzzyklen erfüllt ist, dann kann auf diese Weise auch die Anregungsbedingung erfüllt sein. Das rückgekoppelte Signal wird dann sehr viel größer als das Eingangssignal, so daß die Bedingung $x(t) \equiv 0$ d.h. $|x(t)| << |y(nT)|$ für bestimmte Zeitabschnitte annähernd erfüllt ist. Wenn der betreffende angeregte Grenzzyklus z.B. aufgrund eines Adaptionsalgorithmus der Quantisierungsstufe eine große Leistung besitzt, dann spielt ein Eingangssignal mit vergleichsweise kleiner Leistung keine Rolle mehr und der Grenzzyklus bleibt bestehen, wenn sich die Existenzbedingung nicht ändert.

Wenn nach Abschnitt 4.2 der Prädiktor aber mit Hilfe des rekonstruierten Signals berechnet wird, dann besitzt er von vorneherein eine kleinere Verstärkung und die Anregungsbedingung für Grenzzyklen ist nicht so leicht gegeben.

Daraus ergibt sich aber, daß der adaptive Prädiktor nicht aus dem Originalsignal berechnet werden darf, wenn im Sender und Empfänger eines solchen Quellcodierungssystems rekonstruierte Signalwerte zur Prädiktion benutzt werden und eine grobe Quantisierung betrachtet wird. Bei einem rückgekoppelten System ist es dann das Sinnvollste, den Prädiktor aus dem rückgekoppelten Signal selber zu berechnen.

Wenn gewährleistet ist, daß keine Grenzzyklen entstehen können bzw. wenn im Experiment beobachtet wird, daß Grenzzyklen nicht auftreten, dann hängt ein optimales Verhalten des hier betrachteten Deltamodulators nur noch von der Adaption der Quantisierungsstufe, d.h. vom Adaptionsalgorithmus und von der Art der Schätzung der Reflexionskoeffizienten ab. Die Optimierung während des Codiervorganges selber ist dann bei einem solchen Deltamodulator, genau wie bei einem Deltamodulator mit festem Rekonstruktionsnetzwerk, nur noch von den 1-bit-Entscheidungen zu jedem Abtastzeitpunkt abhängig.

Bessere Ergebnisse als mit einem solchen Deltamodulator mit einer entsprechenden Anzahl von Filterstufen können dann nur noch z.B. mit einer optimalen nichtrückgekoppelten Baumcodierung erreicht werden bzw. mit einer Optimierung der DM nicht zu jedem Abtastzeitpunkt, sondern über ein beliebig langes Sprachsegment (delayd encoding).

6. Optimierung und Beurteilung der Rekonstruktionsgüte von synthetischen Sprachsignalen

Die Beurteilung der Qualität des rekonstruierten Signals bei der Sprachcodierung erweist sich als sehr schwierig. Objektive Gütemeßverfahren, wie die Bestimmung des SNR für lange Signalabschnitte oder Mittelwerte des SNR für eine Anzahl von kurzen Segmenten, geben nicht unbedingt die subjektiv empfundene Qualität wieder, besonders wenn man verschiedene Quellcodierungsverfahren vergleicht |46|. Es sind allerdings für Telefonverbindungen nach CCITT bestimmte minimale SNR-Werte für digitale Modulationssystem festgelegt und zwar für bestimmte Testsignale†. Bei den hier untersuchten niedrigen Übertragungsbitraten sind diese SNR-Werte jedoch nicht zu erreichen. Trotzdem ist die subjektiv festgestellte Qualität recht gut. Um festzustellen, wie gut diese Qualität tatsächlich ist, müßte man umfangreiche Testreihen mit einer ausreichenden Anzahl von Testpersonen durchführen. Es gibt zwar Vorschläge für solche Testverfahren (z.B. |47|), aber es existiert keine verbindliche Normung. Aus diesem Grunde werden zum Abschluß dieser Arbeit zur Unterstützung der in diesem Abschnitt durchgeführten objektiven Güteermittlung auch einfache, informelle subjektive Testreihen durchgeführt. Grundlagen aller Testverfahren sind Sprachproben, die vom Heinrich-Hertz-Institut zur Verfügung gestellt wurden (siehe auch Anhang A2).

6.1 Gaußsches, autoregressives Testsignal

Aufgrund der Ausführungen in Abschnitt 2 und 3 bietet es sich an, als Testsignalquelle eine gaußsche, autoregressive Signalquelle zu benutzen. Für eine solche Signalquelle läßt sich exakt berechnen, welcher SNR-Wert überhaupt erreicht werden kann. Ein solches Testsignal ist überdies sehr realistisch, wenn man die Koeffizienten des Synthesefilters, welches z.B. ein PARCOR-Synthesefilter nach Bild 3.6 sein kann, durch eine PARCOR-Analyse direkt aus dem Sprachsignal gewinnt. Eine gaußsche Rauschquelle läßt sich in einem Digitalrechner ebenfalls

---

†CCITT Empfehlung G.712

sehr leicht simulieren. Das so erzeugte synthetische Sprachsignal ist dann ein stationäres Testsignal, dessen Leistungsspektrum entsprechend dem Grad des Synthesefilters genau dem Leistungsspektrum irgendeines auszuwählenden Sprachlautes entspricht. Diese so erzeugten synthetischen Laute sind allerdings im allgemeinen keine verständlichen Sprachlaute. Nur wenn auch rauschartige Sprachlaute analysiert wurden, dann stimmt das synthetisierte Signal entsprechend dem einfachen Modell der Spracherzeugung mit dem analysierten Sprachlaut überein. Damit steht für einen Deltamodulator mit adaptivem Rekonstruktionsnetzwerk ein adäquates und leicht reproduzierbares Testsignal zur Verfügung.

Um das untersuchte Deltamodulationssystem auf diese Weise testen und optimieren zu können, wurde ein umfangreiches, interaktives Analyse-Synthese Programmsystem mit folgenden Merkmalen entwickelt:

1. Fortlaufende PARCOR-Analyse der Originalsprachsignale mit Darstellung des Zeitverlaufs und des Reflexionskoeffizienten k_1.
2. Ausgabe des Leistungsspektrums mit wählbarer Zeitkonstante T_i (siehe Gleichung (4.11)) zu einem frei wählbaren Abtastzeitpunkt.
3. Vorgabe dieses Leistungsspektrums für das synthetische Testsignal.
4. Berechnung des zu erreichenden minimalen MQF's nach der RDT.
5. Simulation eines Quellcodierungsverfahrens (DISL†, ADPCM und APPDM) für das synthetische Testsignal.
6. Berechnung des SNR-Wertes und zwar für das jeweils betrachtete Codierverfahren für ein Signalsegment wählbarer Länge.
7. Für APPDM:

7.1 Ausgabe der Reflexionskoeffizienten des PARCOR-Synthesefilters im Rückkoppelungszweig des Deltamodulators.

7.2 Ausgabe des (logarithmischen) Leistungsspektrums (siehe

---

†Siehe Anhang A3

Gleichung (2.48)) und zwar für die:

7.2.1 Eingangssignalfolge $\{x_n\}$, d.h. S_{zx} nach (2.52)

7.2.2 Ausgangssignalfolge $\{y_n\}$, des PARCOR-Synthesefilters im Deltamodulator, d.h. S_{zy}. Die Berechnung von S_{zy} wurde so vorgenommen, daß das PARCOR-Synthesefilter im Deltamodulator, dessen Reflexionskoeffizienten sich auf das zu codierende Signal eingestellt hatten, mit dem gleichen Signal angeregt wie das PARCOR-Synthesefilter zur Testsignalerzeugung. Zusätzlich wurde dann zu dem aus diesem Signal ermittelten logarithmischen Leistungsspektrum der Wert $10\log_{10}(\sigma_q^2/\sigma_e^2)$ addiert. Dabei ist σ_e^2 die Leistung der Anregungssignalfolge $\{e_n\}$ und σ_q^2 die Leistung der Folge $\{q_n\}$, die das PARCOR-Synthesefilter im Deltamodulator anregt. Da das Leistungspektrum dieser Folge nicht konstant ist (die Entscheidungen des Deltamodulators sind nach Abschnitt 4.1 nicht unkorreliert), wird also nur der Verlauf des Spektrums der Ausgangssignalfolge $\{y_n\}$ wiedergegeben, der durch das PARCOR-Synthesefilter alleine bestimmt wird.

7.2.3 Ausgangssignalfolge des optimalen Rekonstruktionsnetzwerks nach Abschnitt 2.6.2, d.h. $S_{zy,Opt}$ (Gleichung (2.67)).

Das Test- und Optimierungsverfahren soll im folgenden an zwei Beispielen dargestellt werden. Es wird ein stimmhafter und ein stimmloser (synthetischer) Laut untersucht.

6.1.1 Synthetischer Vokal|ä|

Das Leistungsspektrum dieses synthetischen Lautes zeigt Bild 2.4. In Tabelle 5 sind die maximal möglichen SNR-Werte nach der RDT (zu erreichen durch ein optimales Quellcodierungsverfahren nach Abschnitt 2.6.2) sowie die SNR-Werte, die durch die drei zum Vergleich betrachteten Verfahren erreicht wurden, angegeben. Die mit DISL und APPDM bezeichneten Deltamodulationsverfahren wurden bezüglich der zur Verfügung stehenden Parameter soweit wie möglich zur Erreichung eines maximalen SNR-

Wertes optimiert. Das ADPCM-Verfahren wurde nach |48| simuliert. Dieses ADPCM-System ist ähnlich wie der Deltamodulator in Bild 5.2 aufgebaut. Zur Prädiktion wurde aber nur ein fest eingestellter Koeffizient a_1=0.85 benutzt. Statt des Komparators wurde ein adaptiver Mehrbitquantisierer verwendet.

Die Ermittlung der SNR-Werte erfolgte im Frequenzbereich, um diejenigen Frequenzanteile des Störgeräusches herausfiltern zu können, die nicht in den Übertragungsbereich fielen. Bei der Bestimmung des Fehlersignals x_n-y_n wurde eine konstante Verzögerung und ein konstanter Verstärkungsfaktor des Signals y_n berücksichtigt. Beide Größen wurden solange verändert, bis sich ein maximaler Störabstand ergab.

Tabelle 5 Maximale SNR-Werte in dB für synthetischen Vokal |ä| bei C=16 kbit/s

Ideales System nach der RDT/ reales Modulationsverfahren	Maximaler SNR-Wert in dB
RDF (F=16 kHz) bei	
1 bit/Abtastwert	22.42
APPDM (2)/(2/2/1.18/16 ms)	17.3
APPDM (8)/(2/2/1.18/16 ms)	18.9
DISL (3/3/1.15/2/5)	14.8
RDF (F=8 kHz) bei	
2 bit/Abtastwert	22.86
ADPCM (2 bit)	9.03

Der erreichte SNR-Wert in Bezug auf den maximal erreichbaren SNR-Wert ist beim APPDM-Verfahren recht hoch. Auch der krümmungsadaptive Deltamodulator liefert ein überraschend gutes Ergebnis. Das dürfte allerdings an der zu speziellen Optimierung liegen. Das ADPCM-Verfahren fällt dagegen deutlich ab. Das Ergebnis für das APPDM-Verfahren stimmt übrigens gut mit einem Ergebnis von Bunin |49| überein, der gezeigt hat, daß man in einem DPCM-System mit optimalem Prädiktor für die hier betrachtete Signalklasse einen Störabstand erreichen kann, der um höchstens 4.32 dB niedriger ist als der maximal erreichbare

Störabstand. Als optimal wurde dabei der Prädiktor betrachtet (vergl. (3.9)), der aus dem Originalsignal berechnet werden kann. Das ist für nicht zu grobe Quantisierung auch zulässig. Die Simulationsergebnisse in der vorliegenden Arbeit zeigen, daß man bei grober Quantisierung den Prädiktor aus dem rekonstruierten Signal berechnen sollte. Auch das Erreichen des von Bunin bestimmten SNR-Bereiches ist zusätzlich ein guter Hinweis dafür, daß der auf diese Weise berechnete Prädiktor annähernd optimal arbeitet.

Beim APPDM-Verfahren ist es weiterhin aufschlußreich zu betrachten, wie weit das Originalleistungsspektrum, also S_{zx} bzw. das optimale Rekonstruktionsspektrum $S_{zy,Opt}$ schon durch

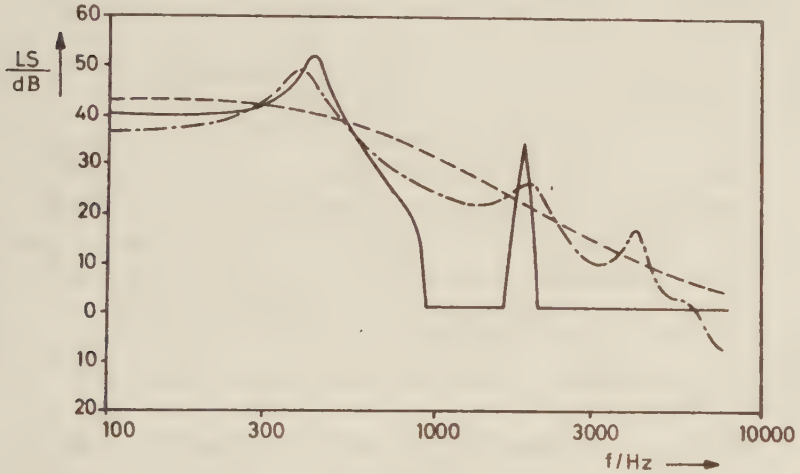


Bild 6.1 Log. Leistungsspektrum LS(f) von:

a) $S_{zy,Opt}$ (———); R=0.9 bit/Abtastwert
Synthetisches, gaußsches Sprachsignal,
Laut: |ä| ; PARCOR-Synthesefilter mit
12 Reflexionskoeffizienten; F=16 kHz

b) $S_{zy} \cdot \sigma_q^2/\sigma_e^2$ (----); $20\log_{10}(\sigma_q/\sigma_e)$= 15.2 dB
für APPDM(2)/(2/2/1.18/16 ms); g_{max}/σ_e=34.6

c) $S_{zy} \cdot \sigma_q^2/\sigma_e^2$ (-•-•-); $20\log_{10}(\sigma_q/\sigma_e)$= 11.9 dB
für APPDM(8)/(2/2/1.18/16 ms); g_{max}/σ_e=34.6

die Reflexionskoeffizienten approximiert wird (siehe Punkt 7.2.2

in Abschnitt 6.1). Bild 6.1 zeigt diese Approximation für zwei verschiedene Rekonstruktionsfiltergrade. Für M=8 ergibt sich schon eine recht gute Annäherung an das optimale Leistungsspektrum. Es wird im wesentlichen der Verlauf des Leistungsspektrums approximiert. Wenn man das PARCOR-Synthesefilter im Deltamodulator (nach stationärer Einstellung der Koeffizienten) mit dem gleichen Eingangssignal wie das PARCOR-Synthesefilter zur Erzeugung des Testsignals anregen würde, dann ergäbe sich ein Ausgangssignal mit geringerer Leistung. Das ist aber bei der Darstellung in Bild 6.1 durch den Faktor σ_q^2/σ_e^2 berücksichtigt. Da beim Deltamodulationsprozeß die zur Verfügung stehende Informationsrate des Codebaums R_{CB} (vergl. (4.3)) nicht voll ausgenutzt werden kann, wurde zur Berechnung von $S_{zy,Opt}$ nur ein Wert von R=0.9 angenommen. Mit diesem Wert wird dann in (2.44) der Wert von θ bestimmt, durch den sich dann mit (2.45) das optimale Leistungsspektrum $S_{zy,Opt}$ bestimmen läßt. Dieses Spektrum entspricht dem dünn ausgezogenen Verlauf von S_{zx} in Bild 2.3 (vergl. auch Bild 2.8).

Da sich die erreichten SNR-Werte für M=2 und M=8 nicht sehr voneinander unterscheiden, bleibt es einem subjektiven Vergleich vorbehalten zu entscheiden, ob sich eine Erhöhung des Rekonstruktionsfiltergrades wesentlich auf den subjektiven Qualitätseindruck auswirkt.

6.1.2 Synthetischer stimmloser Laut |sch|

Das Leistungsspektrum für den synthetischen, stimmlosen Laut ist durch Bild 4.5a gegeben. Die Simulation der verschiedenen Quellcodierungsverfahren erfolgte mit den gleichen Systemparametern wie in Abschnitt 6.1.1. Die erzielten Störabstände zeigt Tabelle 6.

Hier fällt zunächst auf, daß sich bei höheren Rekonstruktionsfiltergraden keine Verbesserung des Störabstandes beim APPDM-Verfahren ergibt. Wenn andere Parameter gewählt werden z.B. ein Adaptionsfaktor k = 1.2, dann wird das System bei M=8 sogar instabil. Die Erklärung dafür dürfte in dem zu hoch gewählten Grad für das Rekonstruktionsnetzwerk liegen. Selbst wenn man bei der Analyse des Originalsignals einen so hohen

Tabelle 6 Maximale SNR-Werte in dB für synthetischen Konsonanten |sch| bei C=16 kbit/s

Ideales System nach der RDT/ reales Modulationsverfahren	Maximaler SNR-Wert in dB
RDT (F=16 kHz) bei	
1 bit/Abtastwert	11.49
APPDM (2)/(2/2/1.18/16 ms)	7.67
APPDM (8)/(2/2/1.18/16 ms)	7.64
DISL (3/3/1.15/2/5)	- 6.3
RDF (F=8 kHz) bei	
2 bit/Abtastwert	13.6
ADPCM (2 bit)	3.2

Grad für rauschartige Signale wählt, erhält man kaum eine bessere Approximation als bei niedrigem Grad.

In Bild 6.2 (Kurve c) ist zu erkennen, daß der Verlauf in der Nähe des Maximums gut approximiert wird. Der maximale Wert des Leistungsspektrums $S_{zy} \cdot \sigma_q^2/\sigma_e^2$ ist allerdings größer als der maximale Wert von $S_{zy,Opt}$. Damit ist eine Anregung von Grenzzyklen unter Umständen möglich. Bei falsch (zu hoch) gewähltem Grad des PARCOR-Synthesefilter gilt, wie festgestellt wurde nicht mehr, daß die im Deltamodulator berechneten Reflexionskoeffizienten aufgrund des Quantisierungsrauschens dem Betrage nach kleiner sind als die Reflexionskoeffizienten des PARCOR-Synthesefilters,welches das Testsignal erzeugt.

Die Approximation des optimalen Rekonstruktionsspektrums ist für M=2 ebenfalls noch recht gut. Bei rauschartigen Signalen sollte im Deltamodulator also ein Synthesefilter mit niedrigem Grad wirksam sein. Wenn man wieder die erreichten SNR-Werte betrachtet, so erkennt man, daß auch bei diesem Signal der maximal erreichbare SNR-Wert vom APPDM-Verfahren um nicht mehr als 4.32 dB unterschritten wird.

Der krümmungsadaptive Deltamodulator wird für dieses Signal sogar instabil. Bei der Optimierung dieses Deltamodulators in Abschnitt 6.1.1, ergaben sich recht große Werte für die Koeffizienten r_1 und r_2 (siehe Bild A3.1) der beiden Integra-

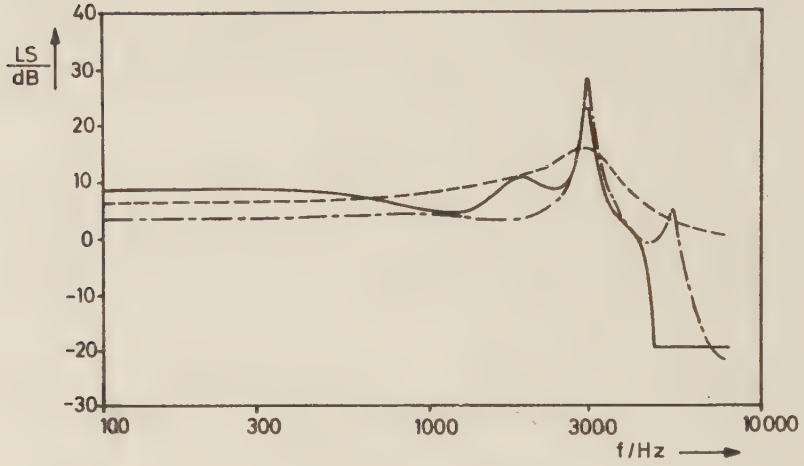


Bild 6.2 Log. Leistungsspektrum LS(f) von:

a) $S_{zy,Opt}$ (———); R=0.9 bit/Abtastwert
Synthetisches, gaußsches Sprachsignal,
Laut: |sch| ; PARCOR-Synthesefilter mit
12 Reflexionskoeffizienten; F=16 kHz

b) $S_{zy} \cdot \sigma_q^2/\sigma_e^2$ (----); $20\log_{10}(\sigma_q/\sigma_e)$=5.53 dB
für APPDM(2)/(2/2/1.18/16 ms); g_{max}/σ_e=44.5

c) $S_{zy} \cdot \sigma_q^2/\sigma_e^2$ (-·-·); $20\log_{10}(\sigma_q/\sigma_e)$=-1.1 dB
für APPDM(8)/(2/2/1.18/16 ms); g_{max}/σ_e=44.5

toren. Um bei diesem Deltamodulator ebenfalls eine Grenzzyklusanregung auszuschließen (insbesondere den 4/03-Mode; siehe Diagramm A3.2), müßten die Koeffizientenwerte der Integratoren verkleinert werden. Dadurch wird der Störabstand für hochkorrelierte Signalsegmente (Vokale) aber auch entsprechend kleiner werden. Das ADPCM-Verfahren arbeitet zwar stabil, erreicht aber nur ungenügende SNR-Werte.

6.2 Frequenzgewichtete Gütekriterien

Zur Erzielung eines möglichst großen Störabstandes ist es nach der RDT notwendig, daß das Spektrum des Fehlersignals (x(t)-y(t)) einen flachen Verlauf erhält (siehe auch Bild 2.3), bis auf Frequenzbereiche, die nicht übertragen werden müssen. Ein solches Fehlersignalspektrum bei der Quellcodierung von Sprachsignalen muß aber nicht notwendigerweise auch subjektiv optimal sein. In |44| wurde gezeigt, daß es für die menschliche Geräuschbewertung günstiger ist, in Bereichen, in denen das Signalleistungsspektrum groß ist, auch eine höhere Geräuschleistung zuzulassen, wenn man nicht in der Lage ist, die Geräuschleistung insgesamt beliebig klein zu machen. Gleichzeitig soll die Geräuschleistung in Bereichen mit kleiner Signalleistung entsprechend abgesenkt werden. Wie das Geräuschspektrum im einzelnen bei unterschiedlichen Lauten verlaufen muß, kann nur durch Versuchsreihen geklärt werden.

Einen weiteren Hinweis auf eine subjektive Geräuschbewertung erhält man aus der objektiven Geräuschmessung bei Telefonverbindungen |50,51,59|. Bei der psophometrischen Geräuschmessung nach CCITT bzw. bei der in den USA gebräuchlichen sogenannten "C-message"-Bewertung wird der Frequenzbereich zwischen 1 kHz und 3 kHz leicht abfallend gewichtet. Geräusche bei hohen Frequenzen fallen also nicht ganz so stark ins Gewicht wie solche bei niedrigen Frequenzen.

Aus diesen Überlegungen heraus wurde von Atal und Schröder |44| folgendes anzustrebendes Geräuschleistungsspektrum G(z) mit A(z) nach Gleichung (2.51) vorgeschlagen:

$$G(z) = K\left|\frac{A(\alpha z)}{A(z)}\right|^2 \qquad (6.1)$$

mit $0 \leq \alpha \leq 1$, $K > 0$.

Mit Hilfe des Parameters α läßt sich dem Geräuschspektrum G(z) eine nahezu beliebige Form im Rahmen der gemachten Überlegungen geben. Für α=1 erhält man ein flaches Geräuschleistungsspektrum und für α=0 hat dieses Spektrum genau den gleichen Verlauf wie das Signalspektrum.

Für ein solches gefordertes Spektrum des Fehlersignals

besteht ebenfalls die Möglichkeit, mit Hilfe der RDT den minimal möglichen Fehler D bei einer bestimmten Informationsrate R zu berechnen |52-54|. Dabei ist der minimal erreichbare mittlere Fehler D am kleinsten für ein flaches Geräuschleistungsspektrum. Die RDF kann man zunächst entsprechend Gleichung (2.41) und (2.42) darstellen:

$$R_N(D) = \frac{1}{N} \sum_{k=1}^{N} \max\left[0, \frac{1}{2} \log\left(\frac{\lambda_k^{(N)}}{\theta_k^{(N)}}\right)\right] \quad , \tag{6.2}$$

$$D(\underline{\theta}) = \frac{1}{N} \sum_{k=1}^{N} \min(\theta_k^{(N)}, \lambda_k^{(N)}) \quad . \tag{6.3}$$

Dabei ist $\underline{\theta} = [\theta_1^{(N)}, \theta_2^{(N)}, \ldots, \theta_N^{(N)}]^T$

der Vektor der Eigenwerte der Kovarianzmatrix des Fehlersignals. Dieser Vektor ist dabei so gewählt, daß sich für die z-Transformierte dieses Vektors das als optimal vorgeschlagene Leistungsspektrum des Fehlersignals nach Gleichung (6.1) ergibt. Bei fester Wahl des Koeffizienten α ist dann der Wert D in (6.3) nur noch von der Konstanten K in (6.1) abhängig, so daß sich eine eindeutige Funktionsdarstellung $R_N(D)$ ergibt.

Mit den gleichen Überlegungen wie in Abschnitt 2.4 gelangt man zu einer Darstellung für den Grenzwert von $R_N(D)$ für $N \to \infty$ und $z = e^{j\varphi}$:

$$R(D) = \frac{1}{4\pi} \int_{-\pi}^{+\pi} \max\left[0, \log\frac{S_{zx}(\varphi)}{G(\varphi)}\right] d\varphi \quad , \tag{6.4}$$

$$D(G[\varphi]) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \min[G(\varphi), S_{zx}(\varphi)] \, d\varphi \quad . \tag{6.5}$$

Diese Betrachtungen können nun dazu dienen, das Ergebnis, besonders von Abschnitt 6.1.1, weiter zu untersuchen. Die subjektive Bewertung des Geräuschspektrums drückt sich durch die Wahl des Faktors α aus. Für die beiden Extremfälle $\alpha=0$ und $\alpha=1$ zeigt Bild 6.3 den Verlauf des Geräuschsspektrums bei maximal erreichbarem Störabstand $SNR_{max}(\alpha)$ bei C=16 kbit/s und F=16 kHz. Für $\alpha=0$ wird nur der untere Frequenzbereich gut genug

übertragen. Für α=1 ist der Störabstand zwar für alle Frequenzen gleich, aber zu niedrig. Das haben Hörversuche in |44| gezeigt. Um eine objektive Bewertung zu erreichen, die dem subjektiven Eindruck möglichst nahe kommt, wird nun ein dritter Wert für α gewählt, so daß für einen Frequenzbereich bis etwa 3.4 kHz gilt (mit f nach Gleichung (2.48)):

$$S_{zx}(f) \geq G(f) \quad . \tag{6.6}$$

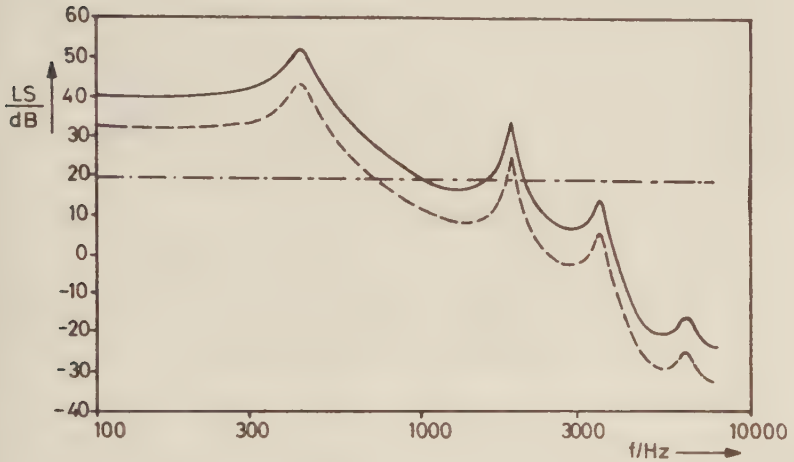


Bild 6.3 Log. Leistungsspektrum LS(f) von:
a) S_{zx} (———); Synthetisches, gaußsches Sprachsignal, Laut: |ä|; PARCOR-Synthesefilter mit 12 Reflexionskoeffizienten; F=16 kHz
b) G (----); $SNR_{max}(\alpha=0.0)$=8.6 dB
c) G (-·-·); $SNR_{max}(\alpha=1.0)$=22.42 dB

Dabei soll mindestens für einen Wert von $f \leq 3.4$ kHz das Gleichheitszeichen in der Bedingung (6.6) gelten. Der Wert für α, der das ermöglicht, kann nur durch ein Iterationsverfahren gefunden werden. Den Verlauf des Geräuschspektrums für diesen Wert von α zeigt Bild 6.4.

Als Vergleich dazu soll jetzt das Geräuschspektrum für das beste Simulationsergebnis für diesen synthetischen Laut betrachtet werden (Bild 6.5). Die Gewichtung für das als ideal

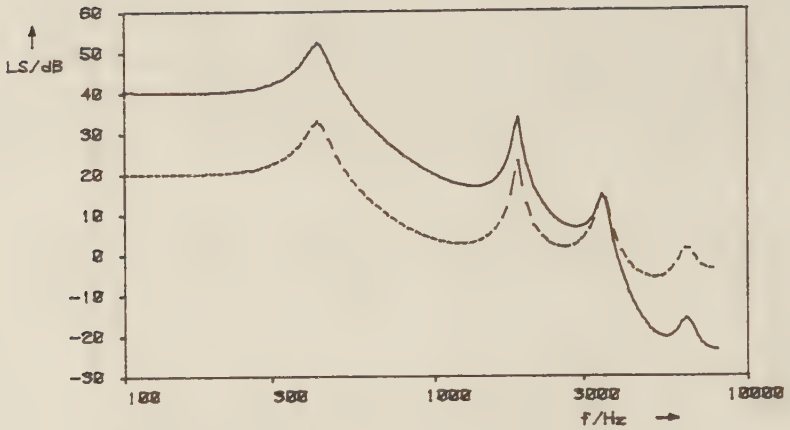


Bild 6.4 Log. Leistungsspektrum LS(f) von:
a) S_{zx} (———); Synthetisches, gaußsches Sprachsignal, Laut: |ä| ; PARCOR-Synthesefilter mit 12 Reflexionskoeffizienten;F=16 kHz
b) G (----); $SNR_{max}(\alpha=0.7)$=19.8 dB

angenommene Geräuschspektrum mit α=0.7 ist im Vergleich dazu ausgeprägter. Für das erreichte Geräuschspektrum in Bild 6.5 ist aber auch eine leichte Gewichtung ab etwa 800 Hz zu erkennen. Der in diesem Fall erreichte Störabstand von 18.9 dB ist nicht mehr sehr weit von dem für α=0.7 erreichbaren Störabstand von 19.8 dB entfernt. Die ersten beiden Maxima (Formanten) im Signalspektrum werden durch das APPDM-Verfahren sogar besser wiedergegeben.

Beim APPDM-Verfahren ergibt sich ein entsprechendes Geräuschspektrum ohne zusätzlichen Aufwand. In |44| wird gezeigt, wie man gezielt ein Geräuschspektrum der Form (6.1) erzeugen kann. Dazu ist ein zusätzliches Netzwerk und die genaue Kenntnis der Abtastwerte des Signals notwendig. Das erfordert aber eine genaue A/D-Wandlung, die beim APPDM-Verfahren ausgeschlossen wurde.

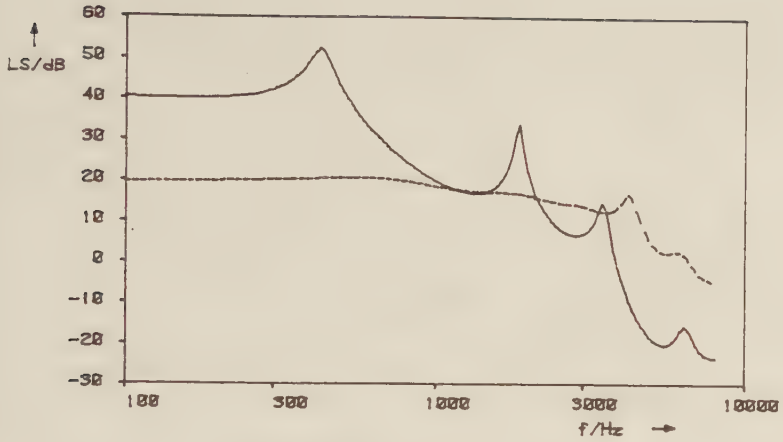


Bild 6.5 Log. Leistungsspektrum LS(f) von:
a) S_{zx}(———); Synthetisches, gaußsches Sprachsignal, Laut: |ä|; PARCOR-Synthesefilter mit 12 Reflexionskoeffizienten;F=16 kHz
b) $\{x_n-y_n\}$(----); SNR=18.9 dB für APPDM(8)/(2/2/1.18/16 ms); g_{max}/σ_e=34.6

6.3 Autoregressives Testsignal mit impulsförmiger Anregung

Wenn man das Testsignal für Vokale noch wirklichkeitsgetreuer gestalten will, so muß man statt der Anregung des PARCOR-Synthesefilters durch eine gaußsche Signalquelle eine Signalquelle mit periodisch sich wiederholenden Impulsen wählen (Abschnitt 3.1). Die Periodenlänge n_pT (Pitch-Periode) ist dabei gleich dem Kehrwert der Sprachgrundfrequenz f_p zu wählen. Ein solches Testsignal wurde in |55| synthetisiert. Angaben über die Form des Anregungssignals bei impulsförmiger Anregung werden z.B. in |60| gemacht. Aufgrund solcher Angaben wurde in |55| die Anregungsfunktion $p(nT)=p_n$ gefunden, die in Bild 6.6 dargestellt ist.

Das mit dieser Impulsfolge und dem PARCOR-Synthesefilter erzeugte Signal x(nT) ist im allgemeinen nicht mittelwertfrei. Bevor dieses Signal codiert wird, wird also noch der Mittel-

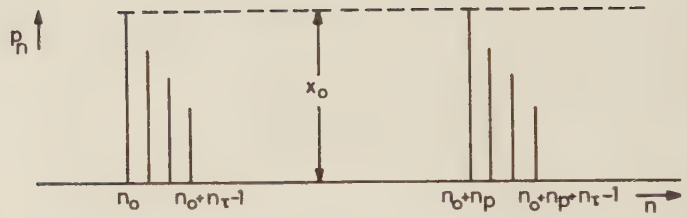


Bild 6.6 Impulsverlauf $p_n = p(nT)$

Impulsbeginn: $n_o = in_p$, $i = 0,1,2,\ldots$

Impulsverlauf $p_n = \begin{cases} x_o e^{-(n-in_p)/n_\tau} & \text{für } in_p \le n \le in_p + n_\tau - 1 \\ 0 & \text{sonst} \end{cases}$

mit $n_\tau \triangleq$ Impulszeitkonstante

wert $\overline{x}$ von jedem Abtastwert x(nT) subtrahiert. Mit der Impulsfolge nach Bild 6.6 wurde jetzt das gleiche PARCOR-Synthese-

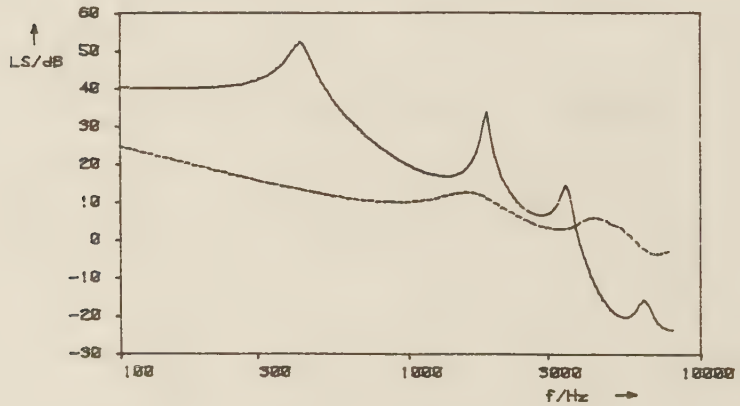


Bild 6.7 Log. Leistungsspektrum LS(f) von:

a) S_{zx}(——); Synthetisches, impulsangeregtes Sprachsignal, Laut: |ä|; PARCOR-Synthesefilter mit 12 Reflexionskoeffizienten;F=16 kHz f_p = 150 Hz; n_τ=4

b) $\{x_n - y_n\}$(----); SNR=19.0 dB für APPDM(4)/(2/2/1.25/16 ms)

filter wie in Abschnitt 6.1.1 (Vokal |ä|) angeregt. Es zeigte sich eine gute Übereinstimmung mit dem Originalvokal |55|.

Das auf diese Weise erzeugte Testsignal wurde nach dem APPDM-Verfahren codiert. Den Verlauf von Signalspektrum, Fehlerspektrum und den erreichten SNR-Wert zeigt Bild 6.7. Da es sich nicht um ein gaußsches Signal handelt, ist hier mit einem größeren zu erreichenden SNR-Wert zu rechnen als für das Signal nach Abschnitt 6.1.1. Der erreichte SNR-Wert von 19.0 dB ist aber nur unerheblich größer als beim gaußschen Testsignal (18.9 dB, Tabelle 5). Ein wichtiges Ergebnis ist aber jetzt, daß das Spektrum des Fehlersignals im ganzen zu übertragenden Frequenzbereich unterhalb des Signalspektrums liegt. Der erste Formant wird jetzt sogar mit ca. 40 dB Störabstand codiert. Beim zweiten Formanten steigt das Spektrum des Fehlersignals leicht an, so wie es beim optimalen Verlauf nach Abschnitt 6.2 für das subjektive Qualitätsempfinden zulässig ist.

Dieser Verlauf des Fehelrsignalsspektrums ist indirekt durch die Güte des Codierverfahrens zu erklären, wenn man annimmt, daß das Spektrum des Quantisierungsfehlers des Differenzsignals $x(nT)+\hat{x}(nT)$ annähernd weiß ist. Der erste Formant wird durch das Rekonstruktionsnetzwerk sehr gut approximiert. In Bild 6.8 ist dazu wiederum der Vergleich von $S_{zy,Opt}$ und $S_{zy} \cdot \sigma_q^2/\sigma_e^2$ (mit $e_n=p_n$) durchgeführt. Aus diesem Grunde ist auch das Fehlerspektrum in diesem Bereich annähernd weiß. Die beiden anderen Formanten werden nur im Mittel approximiert. Zu hohen Frequenzen hin wird auch das Abfallen des Signalspektrums gut approximiert, deshalb fällt auch das Spektrum des Fehlersignals ab.

Insgesamt ist zu erkennen, daß bei diesem Testsignal die Rekonstruktionsqualität (subjektiv) wahrscheinlich größer sein wird als beim gaußschen Testsignal, obwohl der Störabstand in beiden Fällen in etwa gleich ist.

Weiter ist noch zu sagen, daß der Adaptionsfaktor k bei diesem Testsignal im Mittel etwas größer gewählt werden muß (etwa 1.25 gegenüber 1.18 bei F=16 kHz). Es wurden auch nur Rekonstruktionsfiltergrade bis M=6 erreicht. Bei M=8 stellten sich Grenzzyklen ein. Das mag teilweise an der jetzt periodischen (künstlichen) Natur des Testsignals liegen.

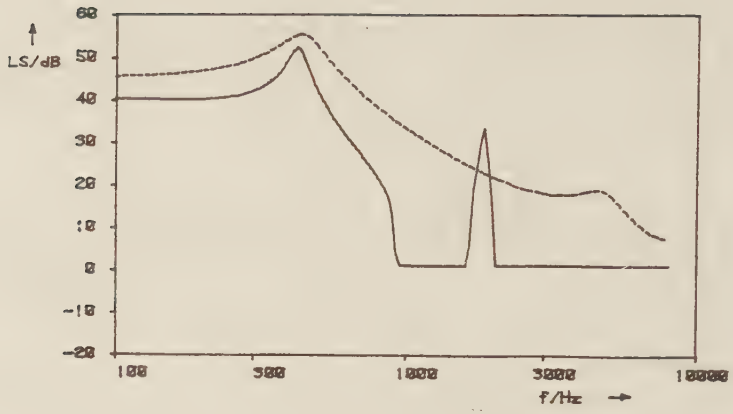


Bild 6.8 Log. Leistungsspektrum LS(f) von:

a) $S_{zy,Opt}$ (———); R=0.9 bit/Abtastwert
Synthetisches, impulsangeregtes Sprachsignal, Laut: |ä|; PARCOR-Synthesefilter mit 12 Reflexionskoeffizienten; F=16 kHz
f_p= 150 Hz; n_τ= 4

b) $S_{zy} \cdot \sigma_q^2/\sigma_e^2$ (----); $20\log_{10}(\sigma_q/\sigma_e)$=20.2 dB
für APPDM(4)/(2/2/1.25/16 ms)

7. Festkommasimulation eines Deltamodulators mit adaptivem PARCOR-Synthesefilter und Augenblicksadaption der Quantisierungsstufe

In Abschnitt 6 wurden alle betrachteten Quellcodierungsverfahren in einer Gleitkommadarstellung simuliert. An dem zur Verfügung stehenden Minirechnersystem (HP 2100) ist es auch möglich, das Ergebnis solcher Simulationsläufe über einen D/A-Wandler auszugeben und damit subjektive Qualitätseindrücke zu gewinnen. Dazu sollten aber Hörproben von mindestens 10 bis 20 s Dauer verarbeitet werden. Da solche Simulationsläufe in Gleitkommarechnung sehr lange Rechenzeiten benötigen (bei APPDM (8) bis zu 4 Std.), wurden Simulationsprogramme in Festkommadarstellung erstellt, die wesentlich kürzere Rechenzeiten benötigen. Gleichzeitig läßt sich dabei untersuchen, welcher Aufwand bei der Realisierung eines solchen Deltamodulators notwendig ist. Die Simulation am Minirechnersystem kann infolge der Mikroprogrammierungsmöglichkeit so realisiert werden, daß dabei schon einige konkrete Hinweise für einen Hardware-Aufbau gefunden werden können. Diese Hinweise betreffen besonders die benötigten Wortlängen der einzelnen Zustandsgrößen und der Reflexionskoeffizienten, aber auch die Art der zu verwendenden Arithmetik.

7.1 Simulation einer PARCOR-Filterstufe

Der wichtigste Teil des untersuchten Deltamodulators ist das adaptive PARCOR-Filter. Aus diesem Grund soll die Simulation einer Filterstufe genauer betrachtet werden. Der aufwendigste Teil einer Filterstufe wiederum ist die Berechnung der Reflexionskoeffizienten k_m. Diese Berechnung geschieht zweckmäßigerweise mit dem Verfahren nach Itakura und Saito (Abschnitt 4.4). Da die Wortlänge des Minirechnersystems, die 16 bit beträgt, nicht überschritten werden sollte, besteht dadurch eine gewisse Beschränkung für die Signalwortlänge innerhalb des Deltamodulators. Außerdem wird so weit wie möglich die Zweierkomplement-Arithmetik des Minirechners benutzt. Die Wortlängen aller Rechengrößen innerhalb des Deltamodulators zeigt Tabelle 7 (vergl. auch

Tabelle 7 Wortlängen der Zustandsgrößen und Koeffizienten bei Simulation auf einem Minirechner

Zustandsgröße, Koeffizient	Registerwortlänge in bit
$y^+_{n,m}$, $y^-_{n,m}$ im Filter	16
$y^+_{n,m}$, $y^-_{n,m}$ im Schätzalgorithmus	15
$(y^+_{n,m} + y^-_{n,m})^2$	32
$(y^+_{n,m} - y^-_{n,m})^2$	32
$f_{n,m}$, $g_{n,m}$	48
$f_{n,m}$ bei Division	32
$g_{n,m}$ bei Division	16
κ_{m+1}	bis zu 16

Bild 4.7).

Zu Beginn der Berechnung eines Reflexionskoeffizienten ist eine Addition und eine Subtraktion notwendig. Damit in der ersten Filterstufe bei der Berechnung von k_1 bei der Addition zweier gleich großer Zahlen, die die volle 16 bit Wortlänge einnehmen können, kein Überlauf eintritt, werden die beiden Größen $y^+_{n,m}$ und $y^-_{n,m}$ jeweils um eine Binärstelle nach rechts verschoben. Das Ergebnis der Filterung der quadrierten Größen wird in zwei 48 bit Registern $f_{n,m}$ und $g_{n,m}$ gespeichert. Um den Divisionsalgorithmus des Minirechners benutzen zu können, werden $f_{n,m}$ auf

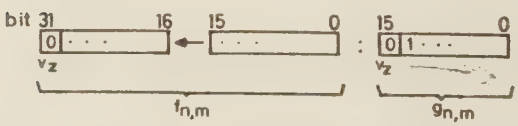


Bild 7.1 Division der Beträge von $f_{n,m}$ und $g_{n,m}$

32 bit und $g_{n,m}$ auf 16 bit beschränkt. Dazu wird der Binärwert in $g_{n,m}$ so weit nach links verschoben, daß das höchstwertige, von Null verschiedene bit von $g_{n,m}$ auf die Stelle 14 eines 16 bit Registers gelangt (siehe Bild 7.1). Die Anzahl der Verschiebungen wird gezählt. Entsprechend dieser Zahl wird auch der Inhalt von $f_{n,m}$ nach links verschoben, nachdem vorher der Betrag dieser Größe gebildet wurde und das Vorzeichen gesondert gespeichert wurde. Um diesen Divisionsalgorithmus mit 48 bit Größen (drei Rechnerworte) durchführen zu können, wurde ein Mikroprogramm entwickelt.

In diesem mikroprogrammierten Divisionsalgorithmus wird nur die gewünschte Anzahl der Quotientenbits erzeugt. Durch diese Art der Division wird k_{m+1} dem Betrage nach immer etwas größer, als es bei der Betrachtung aller zur Verfügung stehenden Binärstellen von $f_{n,m}$ und $g_{n,m}$ der Fall wäre.

Bei den Multiplikationen im PARCOR-Synthesefilter wird die Methode des Betragsschneidens angewendet und zwar so, daß das Multiplikationsergebnis höchstens 16 bit Wortlänge einnimmt.

7.2 Stabilität von Grenzzyklen in Abhängigkeit der Wortlänge der Reflexionskoeffizienten

Obwohl es nicht wünschenswert ist, daß störende Grenzzyklen überhaupt angeregt werden, ist es interessant zu untersuchen, ob einmal angeregte, bestimmte Grenzzyklen in dem betrachteten Deltamodulationssystem bestehen bleiben können. Da sich die Reflexionskoeffizienten auf den Grenzzyklus einstellen, ist es nur notwendig, die Werte der Reflexionskoeffizienten für ein bestimmtes vorgegebenes Bitmuster der b_n, welches einem Mode entspricht, zu berechnen und dann die Existenzmöglichkeit dieses Modes für diese Koeffizientenwerte mit Hilfe des entsprechenden Stabilitätsdiagramms zu überprüfen. Die Werte der zu bestimmenden Reflexionskoeffizienten werden dabei je nach Bitmuster mehr oder weniger von der Wortlänge der Reflexionskoeffizietnen abhängen.

Es soll also ein nicht rückgekoppeltes Deltamodulationssystem (Decodierer) nach Bild 7.2 untersucht werden. Als Eingangssignal wird die für einen bestimmten Mode charakteristische,

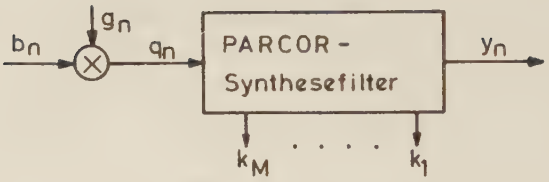


Bild 7.2 Nichtrückgekoppeltes, adaptives PARCOR-Synthesefilter bei eingeprägter, periodischer b_n-Folge

periodische b_n-Folge eingeprägt. Nach einer hinreichend langen Einschwingzeit werden die sich ergebenden Reflexionskoeffizienten betrachtet. Verändert wird bei den einzelnen Moden die Wortlänge der Reflexionskoeffizienten. Es wird nur der stationäre Fall bei der Adaption (Abschnitt 5.2) für die Filtergrade M=2 und M=4 bei der Anregung mit dem

4/03 - Mode,
6/07 - Mode
und dem 6/13 - Mode

untersucht. Der Koeffizient c von W(z) beträgt $1-2^{-8}$ (d.h. T_i=16 ms bei F=16 kHz). Die Quantisierungsstufen g_n werden

Tabelle 8 Reflexionskoeffizienten bei eingeprägten Mode

Mode	k_1	k_2	Koeffizientenwortlänge in bit
4/03	0.0	-0.96875	6
(k=	0.001953	-0.996094	10
1.17)	0.002747	-0.989838	16
6/07	0.40625	-0.90625	6
(k=	0.345703	-0.939453	10
1.17)	0.241943	-0.899811	16
6/13	-0.46875	-0.96875	6
(k=	-0.498047	-0.984375	10
1.25)	-0.498444	-0.999969	16

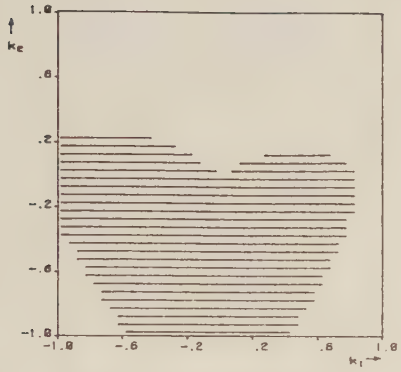


Bild 7.3a
Stabilitaetsbereich
(Koln 4/0003-Mode)
APPDM(2)/(2/2/1.17)

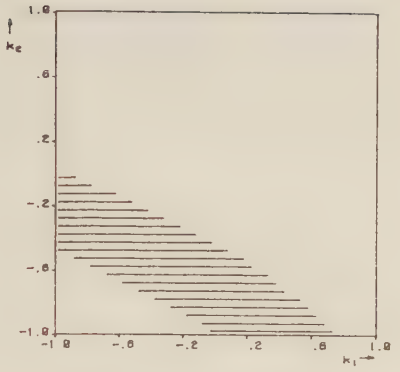


Bild 7.4a
Stabilitaetsbereich
(Koln 4/0003-Mode) k2= -.415078
APPDM(4)/(2/2/1.17) k4= -.608075

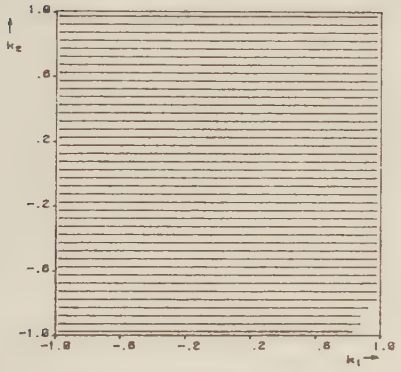


Bild 7.3b
Stabilitaetsbereich
(Koln 6/0007-Mode)
APPDM(2)/(2/2/1.17)

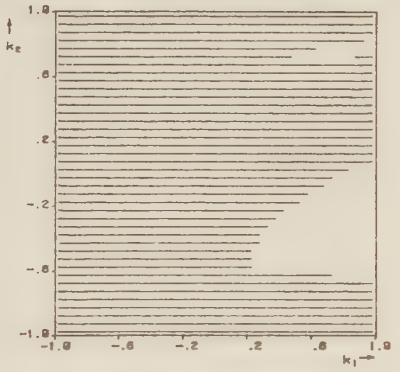


Bild 7.4b
Stabilitaetsbereich
(Koln 6/0013-Mode) k2= .458649
APPDM(4)/(2/2/1.25) k4= -.000060

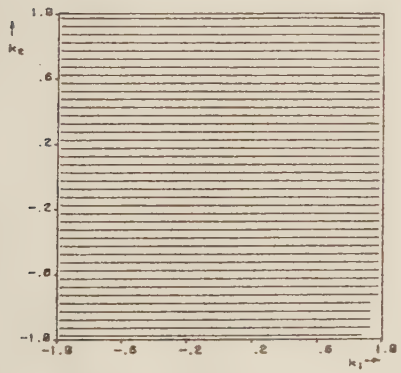


Bild 7.3c
Stabilitaetsbereich
(Koln 6/0013-Mode)
APPDM(2)/(2/2/1.25)

APPDM(4)/(2/2/1.17)

k_2= -0.914917

k_4= 0.160217

6/0007-Mode nicht möglich

entsprechend dem Adaptionsalgorithmus, der von den letzten beiden übertragenen Binärwerten b_n abhängen soll, einem Festwertspeicher entnommen. Da bei dieser Abspeicherung von g_n durch die begrenzte Wortlänge der Speicherinhalte Fehler auftreten, ändert sich der Adaptionsfaktor k je nach dem ob die Adaption gerade bei kleinen oder bei großen Werten von g_n arbeitet.

Da ja die Festkommasimulation in erster Linie zur Ermittlung der subjektiv empfundenen Qualität des entwickelten Deltamodulators dienen soll, wurde im voraus ein Adaptionsfaktor von 1.17 für reale Sprachsignale als annähernd optimal ermittelt (siehe auch Kapitel 8).

Wie man zeigen kann, werden beim 6/13-Mode im stationären Fall aber nur die kleinsten Quantisierungsstufen eingeschaltet, so daß sich aufgrund der Festkommasimulation ein größerer Adaptionsfaktor k=1.25 ergibt.

Für M=2 stellen sich die Werte für k_1 und k_2 nach Tabelle 9 ein. Für alle betrachteten Wortlängen ist aufgrund der Stabilitätsdiagramme in Bild 7.3 zu erkennen, daß keiner der angeregten Moden existieren kann. Wenn einer dieser Moden angeregt wird, dann werden nach kurzer Zeit solche Reflexionskoeffizienten berechnet, die die Existenz dieses Modes nicht mehr ermöglichen.

Für M=4 sind die entsprechenden Ergebnisse in Tabelle 9

Tabelle 9 Reflexionskoeffizienten bei eingeprägtem Mode

Mode	k_1	k_2	k_3	k_4	Koeffizientenwortlänge in bit
4/03	0.0	-0.9375	0.96875	0.1875	6
(k=	-0.636719	-0.75	0.605469	-0.998047	10
1.17)	0.299072	-0.553986	-0.415070	-0.808075	16
6/07	-0.0625	-0.03125	-0.84375	0.3125	6
(k=	0.14451	-0.212891	-0.925781	0.251933	10
1.17)	0.136108	-0.178711	-0.914917	0.160217	16
8/13	-0.375	-0.8125	0.40625	-0.9685	6
(k=	-0.371094	-0.742188	0.373047	-0.984375	10
1.25	-0.458893	-0.897461	0.458648	-0.999969	16

dargestellt. Der 6/07-Mode ist für alle Wertepaare k_3, k_4 im betrachteten k_1, k_2-Bereich nicht existenzfähig. Dieser Mode würde auch sehr stören, da die größten Quantisierungsstufen eingeschaltet würden. Bild 7.4 und im Anhang Bild A 8.1 und A 8.2 zeigen die Stabilitätsdiagramme für die beiden anderen Moden. Anhand dieser Stabilitätsdiagramme ist zu erkennen, daß nur der 4/03-Mode für 10 und 16 bit Koeffizientenwortlänge existenzfähig ist. Bei 6 bit Koeffizientenwortlänge ist dieser Mode aber nicht mehr existenzfähig. Da sich bei diesem Mode zufällige Werte für die Quantisierungsstufen ergeben, sind aber auch andere Stabilitätsdiagramme (anderer Adaptionsfaktor k!) für diesen Mode maßgebend.

Damit scheint es grundsätzlich möglich, die Reflexionskoeffizientenwortlänge so zu wählen, daß bestimmte, einmal angeregte Grenzzyklen nicht bestehen bleiben können.

7.3 Beschränkung des Wertebereiches von k_1 und Änderung der Stufenzahl des PARCOR-Synthesefilters in Abhängigkeit von k_1

Aufgrund eigener Untersuchungen (Abschnitt 5.4.2) und Angaben in der Literatur [21] ist zu sehen, daß die Reflexionskoeffizienten bie Sprachsignalen nicht den vollen Wertebereich benötigen, der bei einem stabilen PARCOR-Synthesefilter erlaubt ist. Insbesondere ist für den Reflexionskoeffizienten k_1 zu erkennen, daß dessen Werte, bei den hier untersuchten Abtastraten, in etwa in dem Bereich

$$0.5 < k_1 < 1.0$$

liegen.

Zu Beginn der Sprachcodierung ist es notwendig, die Anfangswerte der Reflexionskoeffizienten zu definieren. Die sich im Laufe der Codierung zu den Abtastzeitpunkten nT einstellenden Reflexionskoeffizienten k_m sollen jetzt mit $k_{n,m}$, n=0,1,....; m=1,2,...,M bezeichnet werden. Es hat sich bei den Simulationen herausgestellt, daß nur der Anfangswert für k_1, also $k_{0,1}$, ungleich Null gesetzt werden muß. Insbesondere werden deswegen für alle nachfolgend beschriebenen Simulationen

folgende Anfangswerte gesetzt:

$$k_{0,m} = \begin{cases} 0.5 & , m=1 \\ 0 & , m=2,3,\ldots,M \end{cases}$$

Wie sich in Abschnitt 6 gezeigt hat, ist abhängig vom gerade zu codierenden Sprachlaut die Anzahl der benötigten Stufen des Synthesefilters, die eine ausreichende Güte ermöglichen, unterschiedlich groß. Andererseits wächst mit einer größeren Stufenzahl die Möglichkeit der Grenzzyklusanregung besonders in Sprachpausen, da dann die Anregungsbedingung eher erfüllt ist.

Bei energiereichen Sprachsegmenten, also bei Vokalen, ist eine größere Stufenzahl notwendig, um eine gute Anpassung an das Leistungsspektrum zu erreichen. Da die Gefahr der Anregung von Grenzzyklen aber umso mehr abnimmt, je größer und zufälliger das Eingangssignal x(t) ist, ist es sinnvoll die Anzahl der Stufen des PARCOR-Synthesefilters ebenfalls an das Signal anzupassen. Ein gutes Kriterium dafür scheint der Wert des Reflexionskoeffizienten k_1 zu sein. Bei der Codierung von Sprachsignalen werden also bestimmte Schwellen, im weiteren mit Schaltschwellen bezeichnet, für k_1 vorgegeben, bei der eine Anzahl von Stufen des Synthesefilters zu- oder abgeschaltet wird.

Die Einführung einer solchen Kontrolle der Stufenzahl des PARCOR-Synthesefilters hat zwei Vorteile:

1. Bei hochkorrelierten Eingangssignalen ist es möglich, die Qualität des rekonstruierten Signals zu verbessern.
2. Bei schwach korrelierten Eingangssignalen ist eine große Stufenzahl des Synthesefilters nicht notwendig. Gleichzeitig wird die Gefahr der Entstehung von Grenzzyklen stark vermindert. In Sprachpausen ist nur eine Stufe notwendig, so daß dort nur der 2/1-Mode existieren kann (siehe Stabilitätsdiagramme 5.11, 5.12 und A6.1), wenn der Wert von k_1 in etwa <0.7 ist. Das ist aber erfahrungsgemäß in Sprachpausen der Fall. Das Leerlauf-

verhalten des Deltamodulators wird dadurch stark verbessert, da dann aufgrund des verwendeten Adaptionsalgorithmus für die Quantisierungsstufe nur der kleinste Wert g_{min} eingeschaltet wird.

Bei der für die Sprachqualität und die Stabilitätseigenschaften notwendigen variablen Stufenzahl des PARCOR-Synthesefilters zeigt sich ein weiterer Vorteil dieser Filterstruktur gegenüber einer Transversalfilterstruktur. Beim PARCOR-Synthesefilter ist es nur notwenig, die nicht benötigten Koeffizienten zu Null zu setzten. Bei einem Transversalfilter wird bei einer Änderung des Filtergrades ein anderer Koeffizientensatz benutzt, dessen Koeffizientenwerte in anderen Wertebereichen liegen als die des zuvor eingeschalteten Koeffizientensatzes. Bei einer kurzen Koeffizientenwortlänge wären entsprechende Umschaltgeräusche schwer zu vermeiden.

Um gewisse Anhaltspunkte für die Schaltschwellen von k_1 zu erhalten, wird für Sprachsignale die Trajektorie betrachtet, die die Reflexionskoeffizienten k_1 und k_2 für ein Sprachsegment durchlaufen. Gleichzeitig läßt sich dabei auch noch der Einfluß der Reflexionskoeffizientenwortlänge auf das Codierverhalten untersuchen.

Bild 7.5 und 7.6 zeigen zunächst die Codierung des gleichen Sprachsegmentes mit 10 bit und 6 bit Koeffizientenwortlänge. Das rekonstruierte Signal ist nicht gefiltert. Der Kurvenzug entsteht durch eine lineare Verbindung der Abtastwerte. Bei 10 bit Koeffizientenwortlänge ist ein sehr glatter Verlauf zu erkennen. Der maximal erreichte Wert für k_1 beträgt etwa $0.993 \approx 1-2^{-7}$. Um einen so großen Wert erreichen zu können, ist also mindestens eine Wortlänge von 8 bit (mit Vorzeichenbit) erforderlich. Bei einer Wortlänge von 6 bit ergibt sich nur ein maximaler Wert von $0.96875 = 1-2^{-5}$. Dieser Wert reicht aber bei 16 kHz Abtastrate für hochkorrelierte Signale zur Darstellung des Reflexionskoeffizienten k_1 nicht aus. Dabei muß man sich immer verdeutlichen, daß dieser Wert ja aus dem rekonstruierten Signal selber berechnet wird. Bei hochkorrelierten Signalen und einem PARCOR-Synthesefilter vom Grad M=8 zeigen sich also kaum Quantisierungseffekte im rekonstruierten Signal sowie im Reflexionskoeffizienten k_1, sofern die Koeffi-

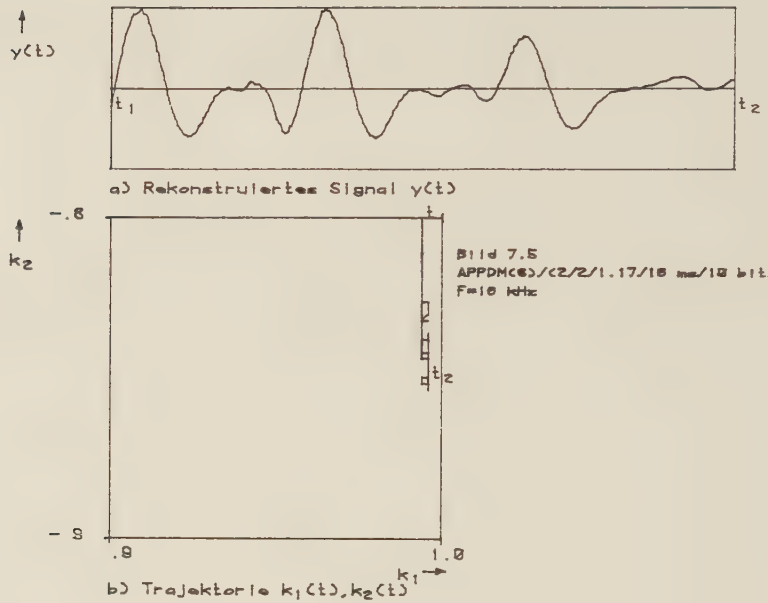


Bild 7.5
APPDM(8)/(2/2/1.17/16 ms/10 bit)
F=10 kHz

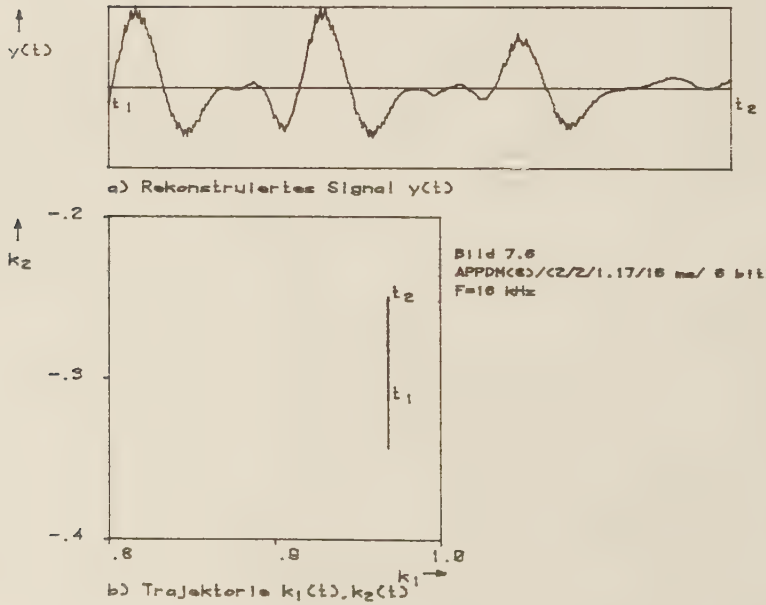


Bild 7.6
APPDM(8)/(2/2/1.17/16 ms/ 8 bit)
F=10 kHz

zientenwortlänge entsprechend groß gewählt wird. Die falsche Wahl der Wortlänge wirkt sich auch auf alle anderen Koeffizienten (in Bild 7.5b und 7.6b ist die Auswirkung auf k_2 zu erkennen) aus. Um die maximal notwendigen Beträge der Reflexionskoeffizienten erreichen zu können (bei F=16 kHz), sind etwa folgende Wortlängen erforderlich:

k_1 , 8 bis 10 bit ,

k_m, m=2,3,...,M , 4 bis 6 bit .

Zur Wirksamkeit der Kontrolle der Anzahl der Stufen des PARCOR-Synthesefilters soll das Einschalten des Deltamodulators während eines energiereichen Sprachsegmentes betrachtet werden. Es wird ein Deltamodulator mit einem Synthesefilter vom Grad M=8 (maximal) untersucht. In Bild 7.7 wird nur der Anfangswert für k_1 vorgegeben:

$$k_{0,1} = 0.5 \cdot$$

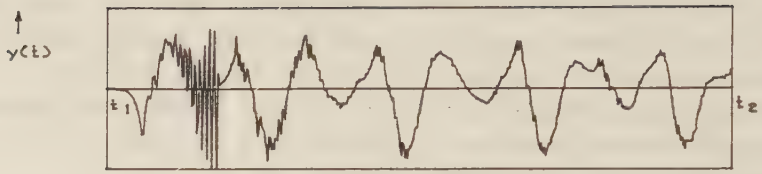
Dieser Wert ist gleichzeitig eine Schaltschwelle s_1 bei der alle anderen Reflexionskoeffizienten außer k_1 zu Null gesetzt werden. Zusätzlich ist dieser Wert auch der minimal erreichbare Wert von k_1. Werden also kleinere Werte als 0.5 für k_1 ausgerechnet, so wird k_1=0.5 gesetzt.

In Bild 7.8 sind dagegen zusätzlich folgende zwei Schaltschwellen eingeführt worden:

$s_{84}=0.96875=1-2^{-5}$, $k_1<s_{84}\rightarrow$ Umschalten von Stufenzahl 8 auf Stufenzahl 4

$s_{42}=0.9375 =1-2^{-4}$, $k_1<s_{42}\rightarrow$ Umschalten von Stufenzahl 4 auf Stufenzahl 2

In Bild 7.7 ist deutlich der Beginn einer Grenzzyklusschwingung zu beobachten. Durch die Schaltschwelle s_1 wird diese Schwingung dann aber wirksam unterbrochen. In den Registern $f_{n,m}$ und $g_{n,m}$ ergeben sich in der Einschaltphase solche Werte, die noch nicht zum gerade zu codierenden Sprachsegment passen. Da-



a) Rekonstruiertes Signal y(t)

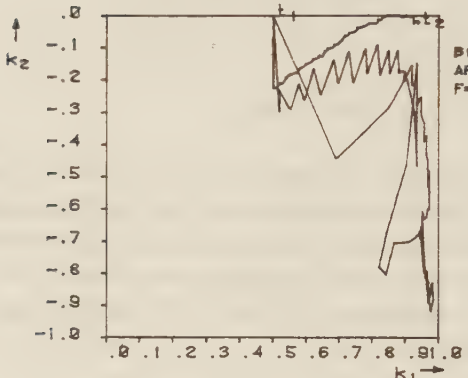
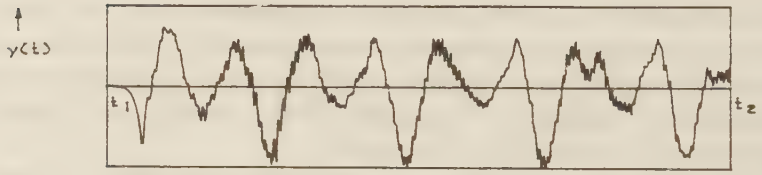


Bild 7.7
APPDM(63)/C2/2/1.17/16 ms/16 bit)
F=16 kHz

b) Trajektorie $k_1(t), k_2(t)$



a) Rekonstruiertes Signal y(t)

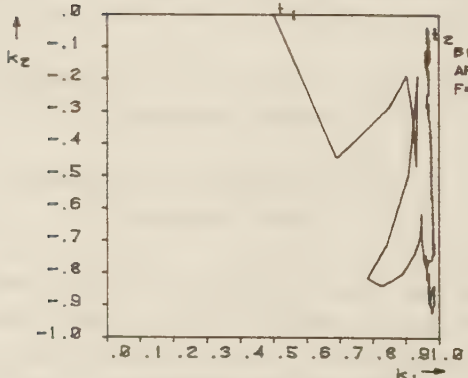


Bild 7.6
APPDM(63)/C2/2/1.17/16 ms/16 bit)
F=16 kHz

b) Trajektorie $k_1(t), k_2(t)$

her ist die Gefahr der Grenzzyklusanregung bei einem so hohen Filtergrad (M=8) gegeben. In Bild 7.8 arbeitet der Deltamodulator in der Einschaltphase nur mit einem 2-stufigen Synthesefilter. Erst nach einer gewissen Einschaltdauer werden je nach Sprachsegment 4 und 8 Stufen des PARCOR-Synthesefilters wirksam. Es findet keine Grenzzyklusanregung statt. Die Einschaltphase dauert aber länger. Wie sich in Versuchen gezeigt hat, kann man die Schaltschwellen für k_1 auch etwas niedriger wählen, so daß häufiger ein Synthesefilter vom Grad M=8 eingeschaltet wird.

7.4 Einfluß von Kanalstörungen

Der Einfluß von Kanalstörungen auf den Gleichlauf von Codierer und Decodierer wurde nicht untersucht. Es sollen an dieser Stelle dazu nur einige Plausibilitätsüberlegungen angestellt werden.

Kanalstörungen, also falsch empfangene binäre Zeichen b_n, bewirken im Decodierer drei verschiedene Arten von Fehlern:

1. Das rekonstruierte Signal y_n wird verfälscht.

2. Die Quantisierungsstufe g_n wird verfälscht.

3. Es werden andere Reflexionskoeffizienten als im Codierer berechnet.

Ohne fehlerkorrigierende Übertragung lassen sich die Fehler der 1. Art nicht vermeiden. Sie haben aber auch die geringste Störwirkung. Die Fehler der 2. Art sind schwerwiegender, besonders bei der betrachteten Augenblicksadaption der Quantisierungsstufe. Diese Fehler werden erst wieder abgebaut, wenn die kleinsten oder größten Quantisierungsstufen hinreichend lange eingeschaltet werden. Um dieses Verhalten zu verbessern, könnte man entweder fehlertolerante Adaptionsalgorithmen wie in [10] benutzen oder die Adaption von einem Mittelwert einer größeren Zahl von binären Zeichen b_n (ca. 10 ms, Silbenadaption) abhängig machen [56]. Eine solche Silbenadaption hat

sich auch als fehlertolerant erwiesen.

Bei der 3. Fehlerart kann man folgende Überlegungen anstellen. Ein falsch übertragenes Binärzeichen führt zu einem falschen Eingangswert q_n für das PARCOR-Synthesefilter. Dadurch wird der Reflexionskoeffizient mit dem größten Index am meisten betroffen. Alle weiteren Koeffizienten werden entsprechend weniger gestört. Aufgrund der Zeitkonstante T_i des Gewichtungsfilters W(z) wird die Auswirkung eines solchen Fehlers entsprechend schnell abgebaut.

Bei gestörten Kanälen sollte man also versuchen, diese Zeitkonstante groß genug zu wählen, damit ein Fehler bei der gewichteten Summierung zur Berechnung der Reflexionskoeffizienten keinen großen Einfluß hat. Andererseits sollte die Zeitkonstante aber auch klein genug sein, so daß ein einmal aufgetretener Fehler schnell genug abklingt.

Eine weitere Verbesserung auf Kosten eines etwas erhöhten Quantisierungsgeräusches ist durch die Wahl einer möglichst großen kleinsten Quantisierungsstufe g_{min} möglich. Dadurch wird in Sprachpausen häufig die Schaltschwelle s_1 für das Abschalten aller Stufen des Synthesefilters, bis auf die erste, erreicht. Auf diese Weise können Codierer und Decodierer in den Sprachpausen wieder synchronisiert werden.

Durch diese Maßnahmen und die in den vorhergehenden Absätzen erwähnten Verfahren zur Unterdrückung der Auswirkung von Kanalstörungen auf den Adaptionsalgorithmus der Quantisierungsstufe sollte es möglich sein, das APPDM-Verfahren auch bei Kanalstörungen zu beherrschen.

8. Ermittlung der subjektiv empfundenen Sprachqualität

Um die tatsächlich erreichbare Qualität des rekonstruierten Signals und das Stabilitätsverhalten des APPDM-Verfahrens im praktischen Betrieb beurteilen zu können, muß auch reales Sprachsignal als Testsignal verwendet werden. Zur Qualitätsbeurteilung wird das Originalsignal mit dem rekonstruierten Signal verschiedener Quellcodierungsverfahren bei unterschiedlichen Kanalkapazitäten von einer hinreichend großen Anzahl von Testpersonen verglichen. Die Beurteilung erfolgt aufgrund einer vorher vereinbarten Güteskala. Die Güteskala und ein Quellcodierungsverfahren wurden genauso gewählt wie in |46|, um einen gewissen Vergleich mit den dort gefundenen Ergebnissen vornehmen zu können. Es wurde dort unter anderem die Güte von ADPCM-Systemen untersucht mit und ohne adaptiven Prädiktor. In der vorliegenden Arbeit wird, zum Vergleich mit dem APPDM-Verfahren, ADPCM und krümmungsadaptive DM untersucht.

Beim APPDM-Verfahren gibt es eine große Anzahl von Systemparametern, die bei den Hörversuchen verändert werden können. Um den Aufwand nicht zu groß werden zu lassen, wurde in vorangehenden Versuchen eine Vorauswahl der zu verändernden Parameter vorgenommen. Dabei stellte sich heraus, daß der Einfluß der Integrationszeitkonstante T_i und des Quantisierungsstufenfaktors k recht gering war, wenn nur annähernd optimale Werte gewählt wurden. Für die Entscheidungsbitmusterlänge des Adaptionsalgorithmus der Quantisierungsstufe wurde ebenfalls schon im voraus ein fester Wert von zwei vorangehenden Binärentscheidungen als annähernd optimal gefunden. In den Hörversuchen wird also nur noch der Einfluß von zwei Parametern untersucht:

1. Anzahl der Stufen des PARCOR-Synthesefilters,
2. Wortlänge der Reflexionskoeffizienten.

Beim krümmungsadaptiven Deltamodulator waren zur Sicherstellung der Stabilität dem Betrage nach kleinere Rückkoppelungskoeffizienten notwendig als die, die sich bei der Optimierung mit synthetischem Sprachsignal ergeben hatten (siehe auch Abschnitt 6.1.1). Besonders der Koeffizient des ersten Integrators mußte stark verkleinert werden.

Alle untersuchten Verfahren wurden auf dem erwähnten Minirechner HP 2100 simuliert. Die Ergebnisse der Simulationsläufe wurden digital auf Magnetband gespeichert und für die Hörversuche über einen 12-bit D/A-Wandler und ein Tiefpaßfilter (Bild A 2.1) auf eine Kompaktkassette ausgegeben.

Neben der hauptsächlich betrachteten Abtastrate von 16 kHz bei der DM wurden auch die beiden Abtastraten 9.6 kHz und 32 kHz untersucht. Dabei wurde bei 9.6 kHz eine Begrenzung des Sprachfrequenzbereichs auf 2.4 kHz vorgenommen.

Die Simulation des APPDM-Verfahrens auf dem Minirechner wurde im wesentlichen entsprechend Bild 4.1 durchgeführt. Als Originalsignale wurden allerdings die vorher entsprechend Anhang A2 abgetasteten und gespeicherten Sprachsignale benutzt. Aus diesem Grunde waren bei den Simulationen, die nicht in Echtzeit durchgeführt werden konnten, der Komparator und der D/A-Wandler nicht notwendig.

Bei den Deltamodulationsverfahren, die ja nicht mit der Nyquistrate arbeiten, ist es jetzt noch möglich, das Quantisierungsgeräusch von der Bandgrenze bis zur halben Abtastrate mit einem festen Digitalfilter zu unterdrücken. Dazu wurde ein Wellendigitalfilter 7. Grades [57] mit kurzer Koeffizientenwortlänge (4 bit) benutzt. Dieses Filter wird in Bild 4.1 zwischen Rekonstruktionsnetzwerk und D/A-Wandler geschaltet. Eine $\frac{\sin(x)}{x}$ -Verzerrung des Ausgangsspektrums nach der D/A-Wandlung wurde nicht korrigiert, da bei der DM die Abtastrate entsprechend groß ist, so daß diese Verzerrung keine große Auswirkung besitzt.

Durch das Wellendigitalfilter wurde das Quantisierungsgeräusch schon um einen hörbaren Anteil vermindert. Es scheint möglich, den Aufwand für dieses Digitalfilter noch kleiner zu halten (z.B. Wellendigitalfilter 5. Grades mit noch kürzerer Koeffizientenwortlänge). Damit bleibt der Aufwand für das Glättungsfilter bei der DM vergleichsweise niedrig.

Bei der ADPCM ist aber immer noch eine recht aufwendige Analogfilterung notwendig. Bei den Hörversuchen wurde aber bei allen Verfahren das gleiche einfache Glättungsfilter benutzt. Deshalb erscheint die ADPCM bei diesen Versuchen möglicherweise geringfügig schlechter, worauf aber bei der Beurteilung Rücksicht genommen werden sollte.

8.1 Ablauf und Ergebnis des Hörversuchs

Es wurden vier verschiedene Versionen der im Anhang A2 dargestellten Sprachprobe (gesprochen von zwei Frauen und zwei Männern) mit den Quellcodierungsverfahren nach Tabelle 10 codiert. Die rekonstruierten Sprachsignale wurden mit Hilfe einer Bewertungsskala mit 9 Gütestufen (9 = sehr gut, ..., 1 = mangelhaft) von 18 Versuchspersonen bewertet. Die Qualitätsstufe 9 soll zusätzlich dabei bedeuten, daß sich das so bewertete rekonstruierte Signal nicht wahrnehmbar vom Original (12 bit, linear quantisiert) unterscheidet.

Tabelle 10 Quellcodierungsverfahren für Hörversuche

Verfahren	C in kbit/s	Symbol
APPDM(2)/(2/2/1.26/13 ms/ 6 bit)	9.6	(x)
APPDM(4)/(2/2/1.26/13 ms/16 bit)	9.6	□
APPDM(2)/(2/2/1.17/16 ms/ 6 bit)	16	
APPDM(2)/(2/2/1.17/16 ms/16 bit)	16	×
APPDM(4)/(2/2/1.17/16 ms/ 6 bit)	16	
APPDM(4)/(2/2/1.17/16 ms/16 bit)	16	□
APPDM(8)/(2/2/1.17/16 ms/ 6 bit)	16	
APPDM(8)/(2/2/1.17/16 ms/16 bit)	16	◇
APPDM(2)/(2/2/1.17/16 ms/16 bit)	32	×
ADPCM(2 bit)	16	○
ADPCM(3 bit)	24	○
ADPCM(4 bit)	32	○
DISL(2/2/1/4)	16	△

Für jedes Quellcodierungsverfahren wurden auf diese Weise (Darbietung der Signale über Kopfhörer) 4•18=72 Bewertungen vorgenommen. Mit Ausnahme des APPDM-Verfahrens bei C=16 kbit/s und 6 bit Koeffizientenwortlänge zeigt Bild 8.1 die mittlere erreichte Güte der einzelnen Verfahren.

Wie aus diesem Bild hervorgeht, ist mit dem APPDM-Verfahren bei C=16 kbit/s und einem PARCOR-Synthesefilter mit bis zu 8 Stufen nahezu die gleiche mittlere Güte zu erreichen wie mit dem ADPCM-Verfahren bei C=32 kbit/s. Dabei ist aller-

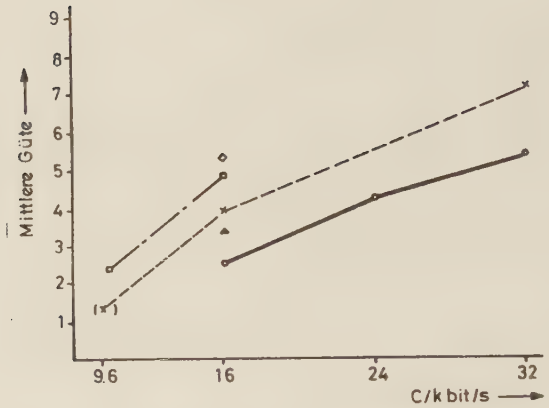


Bild 8.1 Mittlere Güte verschiedener Quellcodierungsverfahren bei realen Sprachsignalen

dings zu bedenken, daß das ADPCM-Verfahren infolge unzureichender Tiefpaßfilterung möglicherweise etwas zu schlecht beurteilt wurde. Berücksichtigt man diese Tatsache, so liegt die zu erreichende Qualtität des APPDM-Verfahrens bei 16 kbit/s sicherlich zwischen der Qualität von 3-bit und 4-bit ADPCM. In [58] wird angegeben, daß Telefonqualität mit 4-bit ADPCM erreicht werden kann.

Das wurde auch von den Versuchspersonen bestätigt, die im Mittel die Telefonqualität mit der Gütestufe 5 bis 6 gleichsetzten. Damit ist mit dem APPDM-Verfahren bei 16 kbit/s Kanalkapazität in etwa Telefonqualität erreichbar. Außerdem ist durch Vergleich mit den Gütewerten in [46] zu erkennen, daß bei dieser Kanalkapazität in etwa die gleiche Güte wie mit der Teilbandcodierung† (subband coding, SBC) und der ADPCM mit adaptivem Prädiktor und zusätzlicher Übertragung der Prädiktorkoeffizienten erreicht wird.

Ein weiteres wichtiges Ergebnis ist, daß das APPDM-Verfahren mit einem PARCOR-Synthesefilter bis zum Grad M=8 für alle Sprachproben ein stabiles Verhalten zeigte. Die Schalt-

† Bei der Teilbandcodierung wird das Spektrum der Sprachsignale in eine Anzahl von Teilbändern aufgeteilt. Jedes Band wird nach Frequenzverschiebung mit adaptiver PCM codiert. Die Teilsignale aller Bänder werden nach Multiplexübertragung einzeln decodiert und summiert.

schwellen entsprechend der Definition in Abschnitt 7.3 betrugen dabei:

$$s_{86} = 0.9375 = 1-2^{-4} ,$$

$$s_{64} = 0.875 = 1-2^{-3} ,$$

$$s_{42} = 0.75 = 1-2^{-2} ,$$

$$s_{1} = 0.5 = 1-2^{-1} ,$$

so daß dieses Verfahren sehr häufig auch mit einem Synthesefilter vom Grad M=8 arbeitete.

Bild 8.2 zeigt die Abhängigkeit der Güte des APPDM-Ver-

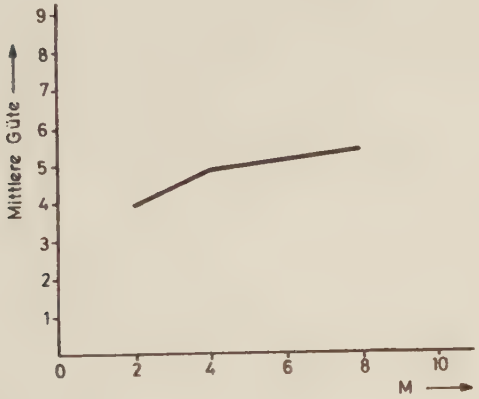


Bild 8.2 Mittlere Güte
APPDM(M)/(2/2/1.17/16 ms/16 bit),
C=16 kbit/s bei realen Sprachsignalen

fahrens vom maximal vorgesehenen Grad M des PARCOR-Synthesefilters. Wie aus diesem Bild zu entnehmen ist, wird sich kaum eine größere Güte ergeben, wenn man bei C=16 kbit/s weit über den Grad M=8 für das Synthesefilter hinausgeht.

Eine Wortlänge von 6 bit für die Reflexionskoeffizienten stellte sich wie erwartet als zu gering heraus. Hierbei wurde bei allen Synthesefiltergraden höchstens eine mittlere Güte von etwa vier erreicht.

Gegenüber der krümmungsadaptiven DM mit festem Rekonstruktionsnetzwerk ergibt sich eine signifikante Verbesserung von etwa zwei Punkten auf der subjektiven Güteskala. Für die Rückkoppelungskoeffizienten der beiden Integratoren wurden die Werte r_1=0.5 und r_2=0.9375 gewählt. Aufgrund des Stabilitätsdiagramms (Bild A 3.2) ist zu erkennen, daß bei dieser Koeffizientenkombination der 4/03-Mode existenzfähig ist. Bei genauer Untersuchung des rekonstruierten Signals stellte sich heraus, daß die Existenzmöglichkeit dieses Modes einen gewissen Einfluß auf den Modulationsprozeß bei den Sprachproben der weiblichen Stimmen hatte. Das zeigt sich auch bei der Gegenüberstellung der erreichten Gütewerte bei den weiblichen und bei den männlichen Stimmen:

Mittlere Güte (weibliche Stimmen) = 3.1 ,
Mittlere Güte (männliche Stimmen) = 3.7 .

Damit ergibt sich ein weiteres wichtiges Ergebnis für Deltamodulatoren mit festem Rekonstruktionsnetzwerk:

> Die Koeffizienten müssen bei 16 kHz Abtastrate so gewählt werden, daß auch der 4/03-Mode nicht existieren kann. Daraus ergibt sich bei der Betrachtung des Stabilitätsdiagramms für diesen Mode, daß praktisch nur ein einfach integrierendes Netzwerk verwendet werden darf, da sich bei der Optimierung gezeigt hat, daß der Koeffizient des zweiten Integrators r_2>0.9 sein sollte.

Wenn die Koeffizienten r_1<0.2 und r_2>0.9 gewählt werden, daß nach dem Stabilitätsdiagramm A 3.2 der 4/03-Mode nicht auftreten kann, dann verschwinden die nichtlinearen Einflüsse beginnender Grenzzyklen, aber das unkorrelierte Quantisierungsgeräusch wird deutlich größer. Dadurch wird die subjektiv empfundene Güte im allgemeinen noch etwas kleiner werden.

Aufgrund dieser Stabilitätsforderung ergibt sich also noch ein weiterer zwingender Grund zur Adaption des Rekonstruktionsnetzwerkes, wenn man höhere Rekonstruktionsgüten bei niedrigen Kanalkapazitäten erreichen will. Die Koeffizienten müssen auf jeden Fall so gewählt werden, daß in Pausen oder

bei abklingenden Signalen keine Grenzzyklen angeregt werden, die nichtlineare Verzerrungen hervorrufen. Dieser Stabilitätsforderung ist aber beim APPDM-Verfahren soweit wie möglich Rechnung getragen.

Bei der Betrachtung der Gütewerte bei C=9.6 kbit/s und C=32 kbit/s läßt sich folgendes sagen:

1. Bei 9.6 kbit/s ist die Sprachqualität nicht ausreichend, auch wenn die Sprache weitgehend verständlich ist.
2. Bei 32 kbit/s wird vom APPDM-Verfahren die höchste Qualität von allen untersuchten Verfahren erreicht. Mit der krümmungsadaptiven DM läßt sich allerdings eine ähnliche Qualität erreichen. Es bleibt noch zu untersuchen, ob sich bei dieser Kanalkapazität eine noch bessere Qualität erreichen läßt, wenn man die Anzahl der Stufen des PARCOR-Synthesefilters und gegebenenfalls die Bandbreite des Sprachsignals (auf etwa 5 kHz) erhöht.

9. Zusammenfassung

Ausgehend von den Ergebnissen der RDT wurde numerisch gezeigt, daß sich bei der Überabtastung von Signalquellen, so wie es bei der DM üblich ist, nicht die optimal mögliche Übertragungsgüte bei fester, vorgegebener Kanalkapazität erreichen läßt. Wenn eine möglichst große Güte erreicht werden soll und eine entsprechend große Kanalkapazität zur Verfügung steht, bedeutet das, daß z.B. die DPCM der DM vorzuziehen ist.

Bei hinreichend niedrigen Kanalkapazitäten (<20 kbit/s) und einer Sprachsignalbandbreite von etwa 3.4 kHz besteht jedoch praktisch kein Unterschied zwischen der DM und der DPCM in Bezug auf die zu erreichende Übertragungsgüte.

Da die DM mit festem Rekonstruktionsnetzwerk bei so niedrigen Kanalkapazitäten keine befriedigende Sprachqualität liefert, wurde versucht die Qualität durch Adaption des Rekonstruktionsnetzwerkes zu verbessern. Grundlage für die Optimierung des Deltamodulationsprozesses sind die Modellvorstellungen zur Spracherzeugung und ein Quellcodierungstheorem für solche Signalmodelle. Aufgrund dieses Theorems läßt sich für eine Baumcodierung das jeweilige optimale Rekonstruktionsnetzwerk bestimmen.

Da durch die DM der einfachste mögliche Baumcode realisiert wird, erhält man gewisse Hinweise zur Adaption des Rekonstruktionsnetzwerkes in einem Deltamodulator. Es stellte sich heraus, daß sich die Anpassung des Deltamodulators an die unterschiedlichen statistischen Eigenschaften des Sprachsignals vorteilhaft mit der Struktur eines PARCOR-Synthesefilters erreichen läßt.

Bevor man überprüfen kann, welche Sprachqualität sich mit einem solchen System erzielen läßt, ist es wichtig die Stabilität des Deltamodulationsprozesses sicherzustellen. Da ein Deltamodulationssystem ein nichtlineares, rückgekoppeltes Abtastsystem darstellt, sind in diesem System unter bestimmten Bedingungen Grenzzyklen möglich. Wenn solche Grenzzyklen auftreten, wird der Deltamodulationsprozeß beeinflußt oder zerstört, auch wenn die Grundfrequenz der Grenzzyklusschwingung oberhalb des Sprachfrequenzbereich liegt und diese Schwingung somit durch ein entsprechendes Filter unterdrückt werden kann. Es wurde all-

gemein gezeigt, wie die Existenzmöglichkeit einer bestimmten Grenzzyklusschwingung exakt nachgewiesen werden kann.

Mit Hilfe dieses Existenznachweises wurden Stabilitätsbereiche (Bereiche der Nichtexistenz) für die wichtigsten Moden in Abhängigkeit der Koeffizienten des Rekonstruktionsnetzwerkes angegeben. Dabei wurde auch die Augenblicksadaption der Quantisierungsstufe berücksichtigt.

Wenn die Existenzbedingung für einen bestimmten Mode erfüllt ist, bedeutet das noch nicht, daß dieser Mode auch unbedingt auftreten muß. Damit ein Mode tatsächlich auftritt, müssen bestimmte Anfangs- und Anregungsbedingungen erfüllt sein. In diesem Zusammenhang stellte es sich als günstig heraus, daß die Reflexionskoeffizienten aus dem rekonstruierten Signal berechnet werden. Dadurch wird die Verstärkung des Rekonstruktionsnetzwerkes nicht zu groß - wenn der Filtergrad richtig gewählt worden ist - und die Anregungsbedingung wird nicht erfüllt.

Um bei Sprachsignalen eine möglichst große Güte zu erreichen und die Stabilität zu gewährleisten, ist es aus diesem Grunde erforderlich, die Anzahl der PARCOR-Filterstufen ebenfalls an das Signal anzupassen. Ein gutes Kriterium für diese Anpassung ist der Wert des Reflexionskoeffizienten k_1.

Die Güte eines solchen Deltamodulators wurde zunächst mit einem synthetischen gaußschen Sprachsignal, für das die optimal erreichbare Qualität vorausberechnet werden kann, überprüft. Dabei stellte sich heraus, daß der optimale Störabstand bis auf 3 bis 4 dB erreicht werden kann. Eine weitere geringfügige Verbesserung ist nur noch mit aufwendigeren Quellcodierungsverfahren möglich.

Bei einem Test mit synthetischen, impulsangeregten Sprachsignalen stellte es sich heraus, daß sogar gewisse subjektive Forderungen an die Gewichtung des Quantisierungsgeräusches eingehalten werden können.

In einem abschließenden subjektiven Hörversuch ergab sich, daß sich mit einem Deltamodulator mit bis zu acht PARCOR-Filterstufen bei C=16 kbit/s annähernd Telefonqualität erreichen läßt. Der Qualitätsunterschied zu einem krümmungsadaptiven Deltamodulator ist recht deutlich, da mit Rücksicht auf die Stabilität dieses Systems bei 16 kbit/s die erste Integrationsstufe kaum

wirksam werden darf. Dadurch ergibt sich ein großes unkorreliertes Quantisierungsgeräusch.

Mein Dank gilt Herrn Dr. F.J. Reich für die Anregung zu dieser Aufgabenstellung und zahlreiche wertvolle Diskussionen sowie Herrn Prof. Dr. A. Fettweis für die Möglichkeit, diese Arbeit am Lehrstuhl für Nachrichtentechnik der Ruhr-Universität Bochum durchführen zu können.

Literaturverzeichnis

1. Deloraine, E.M., Van Mierlo, S., Derjavitch, B., Méthode et système de transmission par impulsions, Französisches Patent Nr. 932140, August 1946.
2. de Jager, F., Delta modulation - a new method of PCM transmission using the 1 - unit code, Philips Res. Rept. No 7, 1952, S. 442 - 466.
3. Winkler, M.R., High information delta modulation, IEEE Int. Conf. Rec. Pt. 8, 1963, S. 260 - 265.
4. Abate, J.E., Linear and adaptive delta modulation, Proc. IEEE No 3, März 1967, S. 298 - 308.
5. Tomazawa, A., Kaneko, H., Companded delta modulation for telephone transmission, IEEE Trans. on Communications, COM-16, Februar 1968, S. 149 - 157.
6. Brolin, S.J., Brown, J.M., Companded delta modulation for telephony, IEEE Trans. on Communications, COM-16, Februar 1968, S. 157 - 162.
7. Greefkes, J.A., de Jager, F., Continuous delta modulation, Philips Res. Rept. No 23, 1968, S. 233 - 246.
8. Jayant, N.S., Adaptive delta modulation with one - bit memory, BSTJ 49, März 1970, S. 321 - 342.
9. Song, C., Garodnik, J., Schilling, D.L., A variable-step-size robust delta modulator, IEEE Trans. on Communications, COM-19, Dezember 1971, S. 1033 - 1044.
10. Schindler, H.R., Entwurf von Deltamodulationsverfahren mit optimalen Übertragungseigenschaften für stationäre und nichtstationäre Signale, Diss. Nr. 4955 ETH Zürich, 1972.
11. van de Weg, H., Quantizing noise of a single integration delta modulation system with an N-digit code, Philips Res. Rept. No 8, 1953, S. 367 - 385.
12. Zetterberg, L.H., A comparison between delta and pulse code modulation, Ericsson Technics 11, Januar 1955, S. 95 - 154.
13. O'Neal, J.B. Jr., Delta modulation quantizing noise analytical and computer simulation results for gaussian and television signals, BSTJ 45, Januar 1966, S. 117 - 141.
14. Goodman, D.J., Delta modulation granular quantization noise, BSTJ 48, Mai-Juni 1969, S. 1197 - 1218.
15. Slepian, D., On delta modulation, BSTJ 51, Dezember 1972, S. 2101 - 2137.

16. Schleifenbaum, B., Informationstheoretische Untersuchungen von adaptiven Deltamodulatoren, Diplomarbeit, Lehrst. f. Nachrichtentechnik, Ruhr-Universität Bochum 1975.
17. Schleifenbaum, B., Informationstheoretische Untersuchung von krümmungsadaptiven Deltamodulatoren, NTZ 29, Oktober 1976, S. 752 - 755.
18. Hübner, H.P., Optimierung des Adaptionsalgorithmus für ein- und zweifach integrierende Delta-Modulatoren, Diplomarbeit, Lehrst. f. Nachrichtentechnik, Ruhr-Universität Bochum, 1977.
19. Berger, T., Rate distortion theory, a mathematical basis for data compression, Prentice-Hall, Inc. Englewood Cliffs, New Jersey, 1971.
20. Flanagan, J.L., Speech analysis synthesis and perception, Springer-Verlag, Berlin, Heidelberg, New York, 1972.
21. Markel, J.D., Gray, A.H., Linear prediction of speech, Springer-Verlag, Berlin, Heidelberg, New York, 1976.
22. Atal, B.S., Schroeder, M.R., Adaptive predictive coding for speech signals, BSTJ 48, Oktober 1970, S. 1973 - 1986.
23. Cohn, D.L., Melsa, J.L., The residual encoder - an improved ADPCM system for speech digitization, IEEE Trans. on Communications, COM-23, September 1975, S. 935 - 941.
24. Noll, P., Adaptive Deltamodulationsverfahren zur Sprachcodierung, Abschlußbericht zum DFG-Forschungsvorhaben, 1975.
25. Cohn, D.L., Melsa, J.L., A new configuration for speech digitization at 9600 bits per second, Proc. ICASSP 79, April 1979, S. 550 - 553.
26. Chang, C.S., An improved residual encoder for speech compression, Proc. ICASSP 79, April 1979, S. 542 - 545.
27. Gelb, A., Vander Velde, W.E., Multiple - input describing functions and nonlinear system design, McGraw-Hill Book Company, New York, St. Louis, San Francisco, Toronto, London, Sydney, 1968.
28. Shannon, C.E., A mathematical theory of communication, BSTJ 27, S. 379 - 423, S. 623 - 656, 1958 (auch als Buch mit Weaver, W., Univ. of Illinois Press, Urbana, Ill., 1949).
29. Schwarz, H.R., Rutishauser, H., Stiefel, E., Numerik symmetrischer Matrizen, B.G. Teubner, Stuttgart, 1968.
30. Grenander, U., Szegö, G., Toeplitz forms and their applications, Univ. of California Press, 1958.

31. Levinson, N., The Wiener RMS (root mean square) error criterion in filter design and prediction, J. Math. Phys., 25, 1947, S. 261 - 278.
32. Fettweis, A., Digital filter structures related to classical filter networks, Arch. Elek. Übertragung, Nr. 25, Februar 1971, S. 79 - 89.
33. Itakura, F., Saito, S., Extraction of speech parameters based upon the statistical method. Proc. Speech Info. Process., Tokohu University, Sendai, Japan, 5.1, 5.12 (in Japanisch, Angabe in |19|), 1971.
34. Haß, C., Digitale Simulation eines adaptiven Deltamodulators, Studienarbeit, Lehrst. f. Nachrichtentechnik, Ruhr-Universität Bochum, 1973.
35. Jacobi, R., Volldigitaler krümmungsadaptiver Delta-Modulator, Diplomarbeit, Lehrst. f. Nachrichtentechnik, Ruhr-Universität Bochum, 1975.
36. Anderson, J.B., Bodie, J.B., Tree encoding of speech, IEEE Trans. on Information Theory, IT-21, Juli 1975, S. 379 - 387.
37. Markhoul, J., Viswanathan, R., Adaptive lattice methods for linear prediction, Proc. ICASSP, Tulsa, 1978, S. 83 - 86.
38. Carter, T.E., Study of an adaptive lattice structure for linear prediction analysis of speech, Proc. ICASSP, Tulsa, 1978, S. 27 - 30.
39. Goldberg, A.J., Schaffer, H.L., A real-time adaptive predictive coder using small computers, IEEE Trans. on Communications, COM-23, Dezember 1975, S. 1443 - 1451.
40. Gersho, A., Stochastic stability of delta modulation, BSTJ 51, April 1972, S. 821 - 841.
41. Steele, R., Delta modulation systems, Pentech Press, London, 1975.
42. Höfelt, M.H.H., On the stability of a 1-bit quantized feedback system, Proc. ICASSP, Washington, April 1979, S. 844 - 848.
43. Noll, P., Sprachübertragung mit adaptiven Differenz-PCM-Verfahren, Technischer Bericht Nr. 164, Heinrich-Hertz-Institut, Berlin-Charlottenburg, 1973.
44. Atal, B.S., Schroeder, M.R., Predictive coding of speech signals and subjective error criteria, IEEE Trans. on Accoustics, Speech, and Signal Processing, ASSP-27, Juni 1979, S. 247 - 254.

45. Katterfeld, M., Untersuchungen zum Aufbau eines APC-Sprachübertragungssystems mit 1-bit-Restfehlerübertragung, 4. DFG-Kolloquium Sprachverarbeitung, Darmstadt, 1. u. 2. Oktober 1979, S. 61 - 64.

46. Tribolet, J.M., Noll, P., McDermott, B.J., Crochiere, R.E., A study of complexity and quality of speech waveform coders, Proc. ICASSP, Tulsa, 1978, S. 586 - 590.

47. Rothauser, E.H., Urbanek, G.E., Pachl, W.P., Isopreference method for speech evaluation, Journal of the accoustical society of America, 1968, S. 408 - 418.

48. Goldstein, L.H., Liu, B., Quantization noise in ADPCM systems, IEEE Trans. on Communications, COM-25, Februar 1977, S. 227 - 238.

49. Bunin, B.J., Rate-distortion functions for gaussian markov processes, BSTJ 48, November 1969, S. 3059 - 3074.

50. Cochran, W.T., Lewinski, D.A., A new measuring set for message circuit noise, BSTJ 39, Juli 1960, S. 911 - 931.

51. Aikens, A.J., Lewinski, D.A., Evaluation of message circuit noise, BSTJ 39, Juli 1960, S. 879 - 910.

52. McDonald, R.A., Schultheiss, P.M., Information rates of gaussian signals under criteria constraining the error spectrum, Proc. IEEE, vol. 52, April 1964, S. 415 - 416.

53. O'Neal, Jr., J.B., Bounds on subjektive performance measures for source encoding systems, IEEE Trans. on Information Theory IT-17, Mai 1971, S. 224 - 231.

54. Davisson, L.D., Rate-distortion theory and application, Proc. IEEE, vol. 60, Juli 1972, S. 800 - 808.

55. Kautz, R.D., Untersuchung eines Deltamodulators mit adaptivem PARCOR-Filter mit synthetischem Sprachsignal, Diplomarbeit, Lehst. f. Nachrichtentechnik, Ruhr-Universität Bochum, 1980.

56. Greefkes, J.A., A digitally controlled codec for speech transmission, IEEE Int. Conf. on Communications, San Francisco 1970, S. 7.33 - 7.47.

57. Wegener, W., Institutsinterne Mitteilungen 1979.

58. Flanagan, J.L., Schroeder, M.R., Atal, B.S., Crochiere, R.E., Jayant, N.S., Tribolet, J.M., Speech coding, IEEE Trans. on Communications, COM-27, April 1979, S. 710 - 737.

59. Rahmig, G., Niederfrequenzübertragungstechnik, Verlag Berliner Union GmbH, Stuttgart, Verlag W. Kohlhammer GmbH, 1972, Stuttgart, Berlin, Köln, Mainz.

60. Rabiner, L.R., Schafer, R.W., Digital Processing of Speech Signals, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1978.

A1 Statistische und informationstheoretische Kenngrößen kontinuierlicher, zeitdiskreter, stationärer Signalquellen (Prozesse)

Signalquelle (Sender):	$\{X, p(\underline{x})\}$
Abtastwerte von x:	$x_n = x(nT)$
Abtastintervall:	T
Signalvektor mit N Abtastwerten (T → transponiert):	$\underline{x} = [x_n, x_{n-1}, \ldots, x_{n-N+1}]^T$
Wahrscheinlichkeitsdichtefunktion (WDF):	$p(x)$ definiert durch die
Wahrscheinlichkeit:	$P(x \leq x_o) = \int_{-\infty}^{x_o} p(x)dx$ (eindimensior
Information:	$\underline{I}(x_o) = \lim_{\Delta x \to o}\{-\log[p(x_o)\Delta x]\}$ $= -\log[p(x_o)] - \lim_{\Delta x \to o}[\log \Delta x]$
Entropie (maßstabsabhängig):	$h(\underline{x}) = -\int_{-\infty}^{\infty} p(\underline{x})\log[p(\underline{x})]\underline{dx}$ (A1.1)
	$(\underline{dx} = dx_n dx_{n-1} \ldots dx_{n-N+1}$, N-faches Integral)
Signalraum (Empfänger):	$\{Y, q(\underline{y})\}$, $\underline{y}$ entspricht $\underline{x}$
Verbundene Signalräume:	$q(\underline{y})$ entspricht $p(\underline{x})$
Verbund-WDF:	$p_v(\underline{x},\underline{y})$ Wahrscheinlichekeitsdic von $\underline{x}$ und $\underline{y}$
Bedingte WDF:	$q_b(\underline{y}/\underline{x}) = \frac{p_v(\underline{x},\underline{y})}{p(x)}$ Wahrscheinlichke dichte von $\underline{y}$ be Kenntnis von $\underline{x}$
	$q_b(\underline{x}/\underline{y}) = \frac{p_v(\underline{x},\underline{y})}{q(y)}$ Wahrscheinlichke dichte von $\underline{x}$ be Kenntnis von $\underline{y}$
Bedingte Entropie:	$h(\underline{y}/\underline{x})$ entsprechend A1.1
	$h(\underline{x}/\underline{y})$ entsprechend A1.1
Transinformation (maßstabsunabhängig):	$T(\underline{x};\underline{y}) = h(\underline{y}) - h(\underline{y}/\underline{x}) = h(\underline{x}) - h($
	Erwartungswert der Information in einem Signalraum bei Kenntnis des zugehörigen (bedingten) Signalvekt im anderen Signalraum. Wenn also z der Signalvektor $\underline{x}$ immer genau ein Signalvektor $\underline{y}$ bedingt, dann ist d Information in $\{Y\}$ über $\{X\}$ vollst dig $(h(\underline{y}/\underline{x}) \equiv 0)$

A2 Beschreibung und Aufbereitung der verwendeten Sprachproben

Zur Gewinnung von Modellparametern aus Sprachsignalen und zur Durchführung von subjektiven Qualitätsuntersuchungen wurden Sprachproben von vier Sprechern (zwei Frauen, zwei Männer) vom Heinrich-Hertz-Institut, Berlin, zur Verfügung gestellt. Die aus diesen Sprachproben für alle Sprecher ausgewählten Sätze lauten:

> "Das Herz Münchens schlägt am Marienplatz. Mögen andere Plätze größer, schöner oder lebendiger sein, der Marienplatz ist der münchnerischste geblieben."
> (ca. 12 s)

Die Sprachproben wurden vor der Abtastung und Quantisierung (12 bit, linear) mit einem aktiven RC-Tiefpaß mit dem Dämpfungsverlauf nach Bild A 2.1 gefiltert.

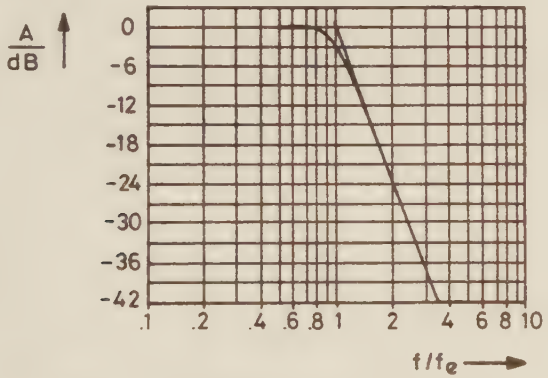


Bild A 2.1 Tiefpaßcharakteristik
Dämpfung A (f/f_e)
f_e = Eckfrequenz (- 3dB)

Die Kurzbezeichnungen (Dateinamen auf Magnetband) der Sprachproben lauten:

Tiefpaßeckfrequenz f_e=3.4 kHz:

MIFM Männlicher Sprecher (I=1,2;F=16,32 [kHz]),
WIFM Weiblicher Sprecher (I=1,2;F=16,32 [kHz]).

Tiefpaßeckfrequenz f_e=2.4 kHz:

MI96M Männlicher Sprecher (I=1,2;F=9.6 kHz),
WI96M Weiblicher Sprecher (I=1,2;F=9.6 kHz).

A3 Krümmungsadaptiver Deltamodulator (DISL) nach [35]

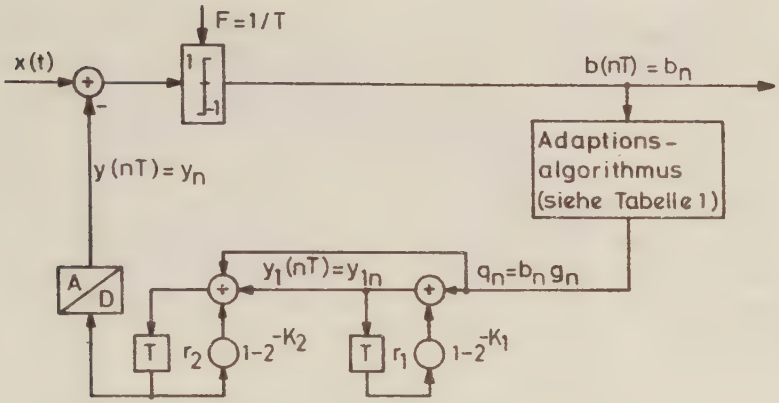


Bild A3.1 Krümmungsadaptiver Deltamodulator.

Kurzbezeichnung für DISL(Deltamodulator mit zweifacher Integration und Sequenzlänge (Bitmusterlänge) zur Adaption der Quantisierungsstufe g_n):

DISL(•/•/•{/•/•})

- K_2 } Angabe kann fehlen
- K_1 }
- Adaptionsfaktor k
- Bitmusterlänge zur Vergrößerung von g_n, $g_n = g_{n-1} \cdot k$
- Bitmusterlänge zur Verkleinerung von g_n, $g_n = g_{n-1} / k$

A3.1 Existenzbedingung für Grenzzyklen beim krümmungsadaptiven Deltamodulator

Für den krümmungsadaptiven Deltamodulator läßt sich mit den gleichen Methoden wie in Abschnitt 5.2 die Existenzbedingung für beliebig vorgebbare Moden überprüfen. Da hier im zweifach integrierenden Rekonstruktionsnetzwerk zwei Zustandsgrößen $y_{1,n}$ und y_n, auftreten, müssen für jeden zu testenden Mode und für jede Koeffizientenkombination r_1, r_2 zwei lineare Gleichungssysteme, deren Ordnung gleich der Periodenlänge L des Modes ist, gelöst werden. Daran anschließend muß die entsprechende Existenzbedingung überprüft werden. Die linearen Gleichungssysteme lauten ($x(t)\equiv 0$):

$$\left.\begin{aligned} y_{1,n+1}-r_1 y_{1,n+1-1} &= b_{n+1} g_{n+1} \\ y_{n+1}-r_2 y_{n+1-1} &= b_{n+1} g_{n+1} + y_{1,n+1} \end{aligned}\right\} \quad l=0,1,\ldots,L-1 \qquad (A3.1)$$

mit der Periodizitätsbedingung

$$y_{1,n+L}=y_{1,n} \quad , \quad y_{n+L}=y_n \qquad \forall\, n.$$

Da b_n von y_{n-1} abhängig ist, erhält man als Existenzbedingung:

$$\mathrm{sgn}(b_{n+1} g_{n+1}) = -\mathrm{sgn}(y_{n+1-1}) \quad , \quad l=0,1,\ldots,L-1 \; . \qquad (A3.2)$$

Für die bei der Codierung von Sprache bei niedrigen Abtastraten optimalen Koeffizientenpaare r_1 und r_2 ist es hauptsächlich notwendig, die Existenz des 4/03-Modes zu überprüfen. Da es sich herausgestellt hat, daß für bestmögliche Rekonstruktionsqualität die Adaption der Quantisierungsstufe von zwei vorangehenden Binärentscheidungen abhängig sein muß, klingen die Zustandsgrößen dieses Modes nicht ab. Obwohl dieser Mode bei 16 kHz Abtastrate eine Grundfrequenz von 4 kHz besitzt und damit bei entsprechender Filterung nicht störend wirkt, kann der Modulationsprozeß bei zufälligen großen Amplituden erheblich gestört werden (nichtlineare Verzerrungen).

Das Stabilitätsdiagramm für diesen Mode und den interessierenden Wertebereich von r_1 und r_2 zeigt Bild A3.2.

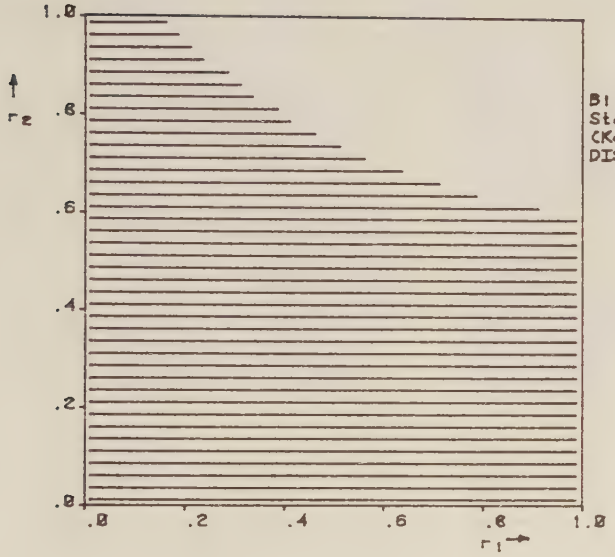


Bild A3.2
Stabilitaetsbereich
(Koln 4/0003-Mode)
DISL(2/2/1.17)

A4 Kurzbezeichnung eines Deltamodulators mit adaptivem PARCOR-Filter und Prädiktion (APPDM)

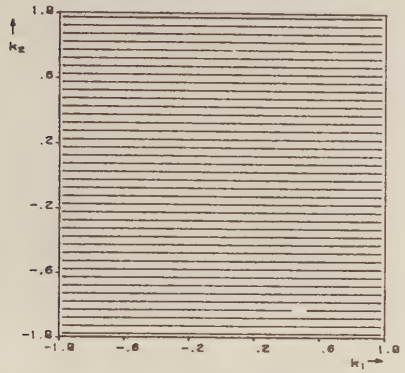
APPDM(·){/(·/·/· {/· ms}{/· bit})}†

- Wortlänge der Reflexions-koeffizienten einschließlich Vorzeichen
- Integrationszeitkonstante T_i
- Adaptionsfaktor k bei Quantisierungsstufe
- Bitmusterlänge der Folge $\{b_n\}$ zur Verkleinerung der Quantisierungsstufe: $g_n = g_{n-1}/k$
- Bitmusterlänge der Folge $\{b_n\}$ zur Vergrößerung der Quantisierungsstufe: $g_n = g_{n-1}{}^k$
- Anzahl M der Stufen des PARCOR-Synthesefilters

† Die geschweiften Klammern deuten an, daß die so eingeklammerten Angaben auch fehlen können; z.B. keine Angabe der Reflexionskoeffizientenwortlänge bei Gleitkommasimulationen.

A5 Stabilitätsbereiche bei 4-stufigem PARCOR-Synthesefilter

(g_n=const.≠0)



k_3= .1
k_4= .1

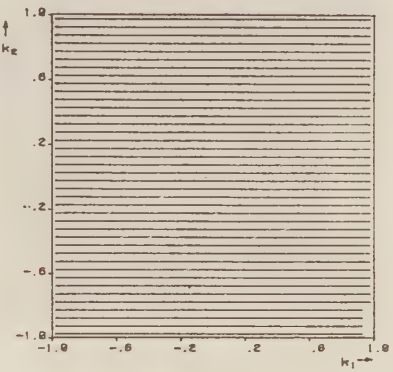


Bild A5.1d
Stabilitätsbereich
(Köln 8/0053-Mode) k_3= -.1
APPDM(4) k_4= -.1

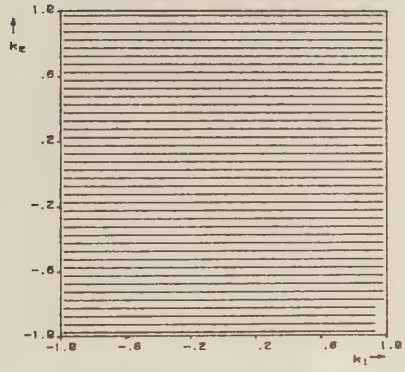


Bild A5.1b
Stabilitätsbereich
(Köln 8/0053-Mode) k_3= .1
APPDM(4) k_4= -.1

APPDM(4)

k_3= 0.2

k_4= 0.2

8/0053-Mode nicht möglich

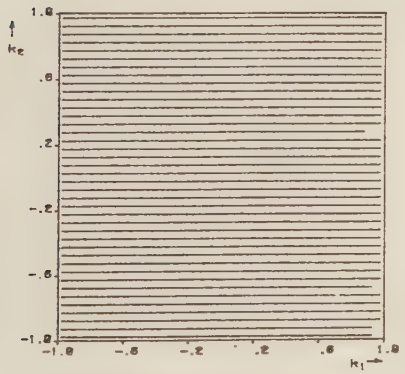


Bild A5.1c
Stabilitätsbereich
(Köln 8/0053-Mode) k_3= -.1
APPDM(4) k_4= .1

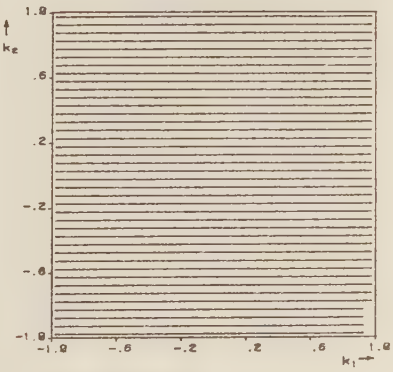


Bild A5.1f
Stabilitätsbereich
(Köln 8/0053-Mode) k_3= .2
APPDM(4) k_4= -.2

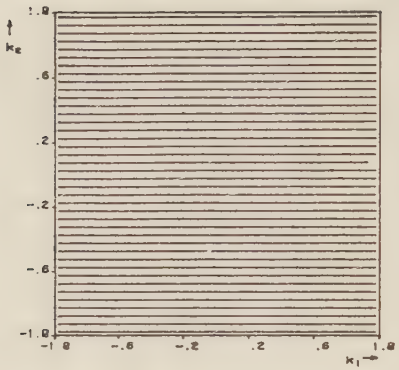


Bild A5.1g
Stabilitätsbereich
(Kein E/0053-Mode) k_3= -.2
APPDM(4) k_4= .2

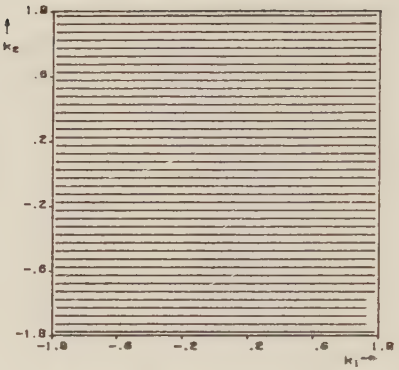


Bild A5.1j
Stabilitätsbereich
(Kein E/0053-Mode) k_3= .4
APPDM(4) k_4= -.4

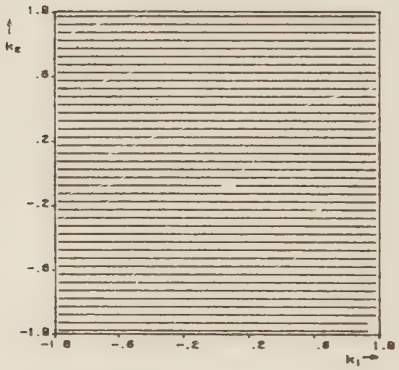


Bild A5.1h
Stabilitätsbereich
(Kein E/0053-Mode) k_3= -.2
APPDM(4) k_4= -.2

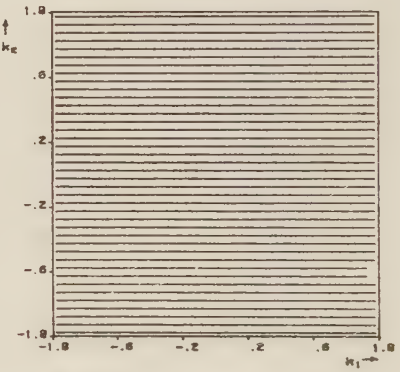


Bild A5.1k
Stabilitätsbereich
(Kein E/0053-Mode) k_3= .5
APPDM(4) k_4= -.5

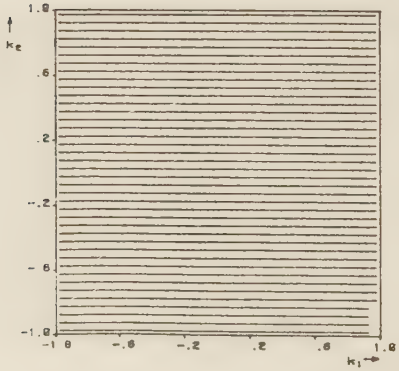


Bild A5.1l
Stabilitätsbereich
(Kein E/0053-Mode) k_3= .3
APPDM(4) k_4= -.3

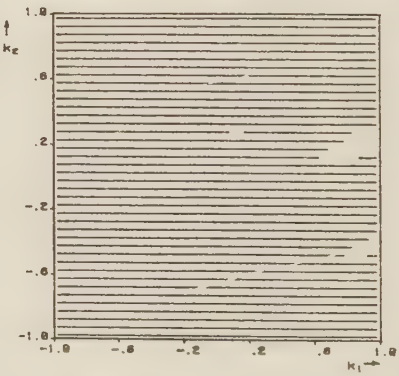


Bild A5.1l
Stabilitätsbereich
(Kein E/0053-Mode) k_3= .6
APPDM(4) k_4= -.6

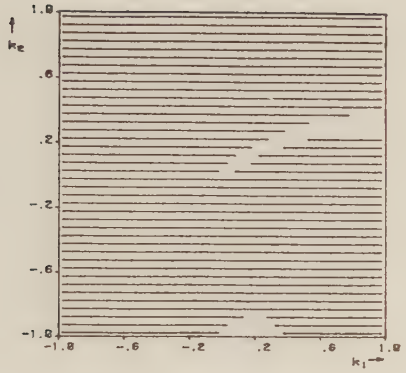


Bild A5.2a
Stabilitätsbereich
(Kein G/BOSS-Mode) k_3= .1
APPDM(4) k_4= .1

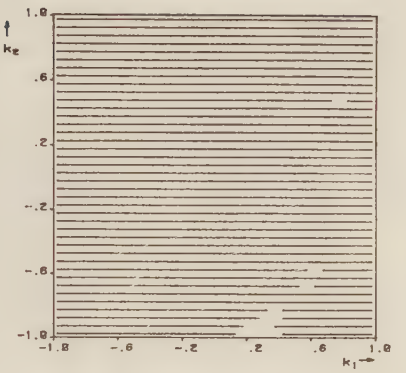


Bild A5.2d
Stabilitätsbereich
(Kein G/BOSS-Mode) k_3= -.1
APPDM(4) k_4= -.1

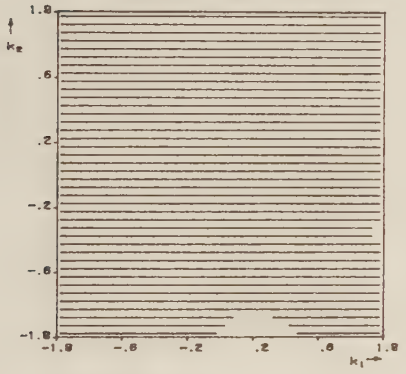


Bild A5.2b
Stabilitätsbereich
(Kein G/BOSS-Mode) k_3= .1
APPDM(4) k_4= -.1

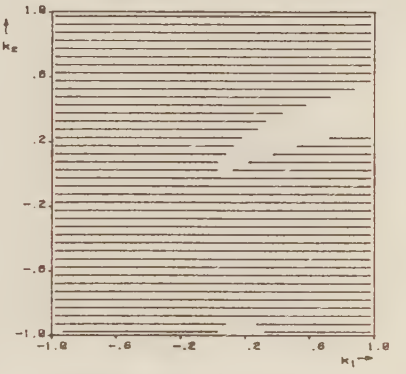


Bild A5.2e
Stabilitätsbereich
(Kein G/BOSS-Mode) k_3= .2
APPDM(4) k_4= .2

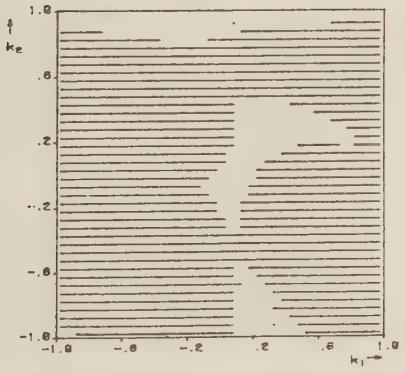


Bild A5.2c
Stabilitätsbereich
(Kein G/BOSS-Mode) k_3= -.1
APPDM(4) k_4= .1

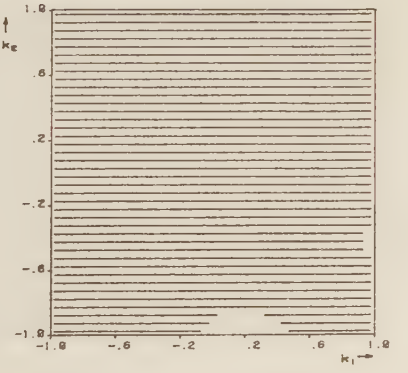


Bild A5.2f
Stabilitätsbereich
(Kein G/BOSS-Mode) k_3= .2
APPDM(4) k_4= -.2

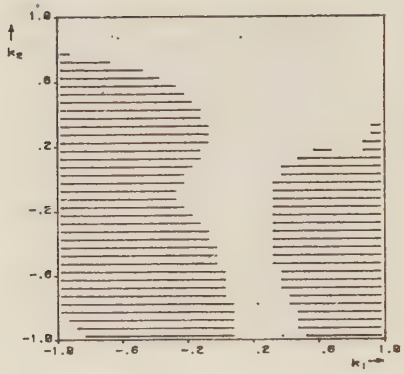


Bild A5.2g
Stabilitätsbereich
(Kein 6/90GS-Mode) k_3= -.2
APPDM(4) k_4= .2

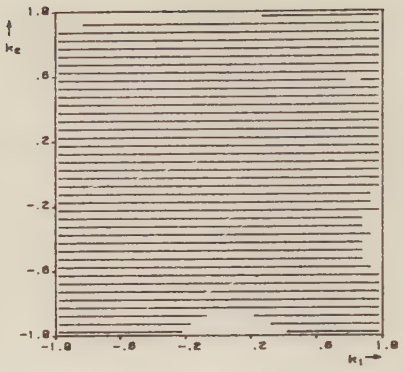


Bild A5.2j
Stabilitätsbereich
(Kein 6/90GS-Mode) k_3= .4
APPDM(4) k_4= -.4

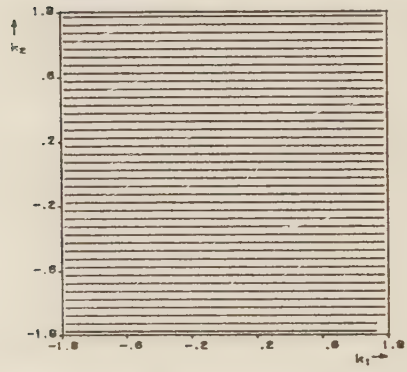


Bild A5.2h
Stabilitätsbereich
(Kein 6/90GS-Mode) k_3= -.2
APPDM(4) k_4= -.2

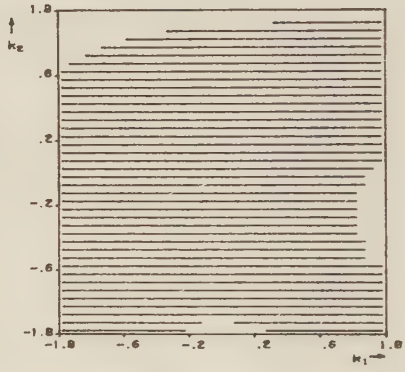


Bild A5.2k
Stabilitätsbereich
(Kein 6/90GS-Mode) k_3= .5
APPDM(4) k_4= -.5

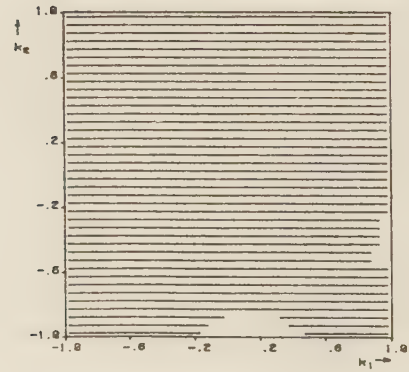


Bild A5.2l
Stabilitätsbereich
(Kein 6/90GS-Mode)
APPDM(4)

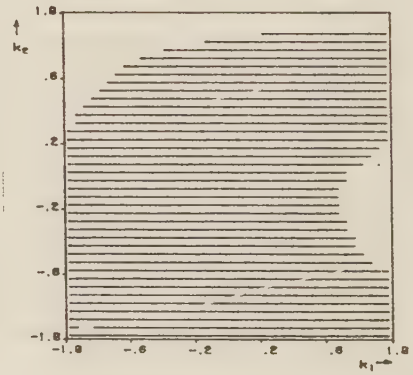


Bild A5.2l
Stabilitätsbereich
(Kein 6/90GS-Mode) k_3= .5
APPDM(4) k_4= -.6

A6 Stabilitätsbereiche bei adaptiver Quantisierungsstufeneinstellung (M=2)

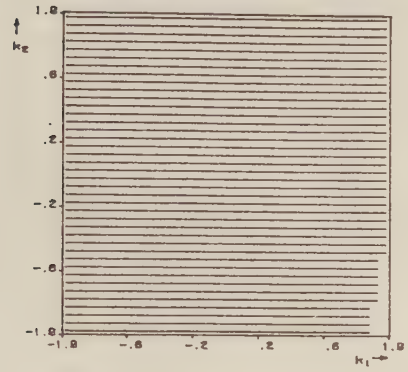


Bild A6.1a
Stabilitätsbereich
(Kein 6/8013-Mode)
APPDM(2)/(2/2/1.10)

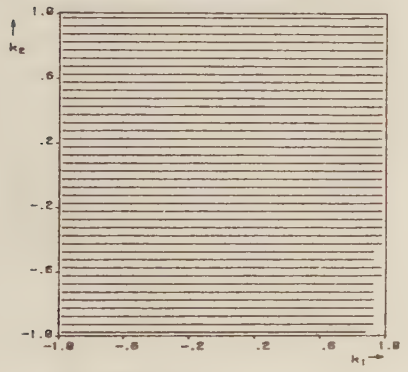


Bild A6.1b
Stabilitätsbereich
(Kein 6/8013-Mode)
APPDM(2)/(2/2/1.20)

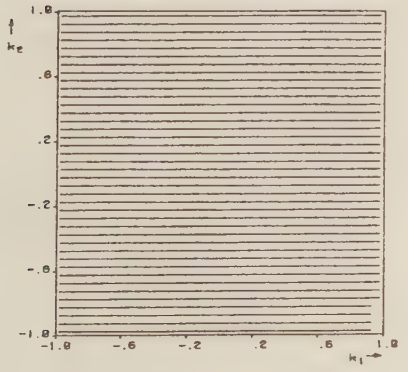


Bild A6.1c
Stabilitätsbereich
(Kein 6/8013-Mode)
APPDM(2)/(2/2/1.30)

A7 Stabilitätsbereiche bei adaptiver Quantisierungsstufeneinstellung (M=4)

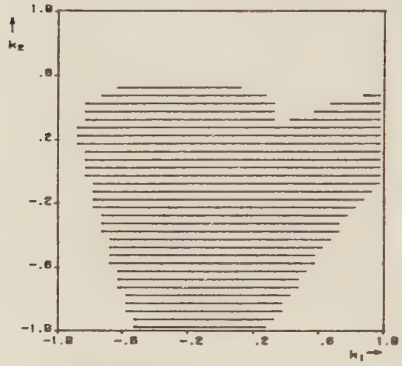


Bild A7.1a
Stabilitätsbereich
(Kein 4/8883-Mode) k_3= .2
APPDM(4)/(2/2/1.28) k_4= .2

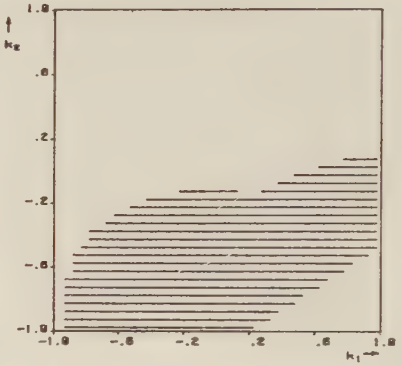


Bild A7.1d
Stabilitätsbereich
(Kein 4/8883-Mode) k_3= .3
APPDM(4)/(2/2/1.28) k_4= -.3

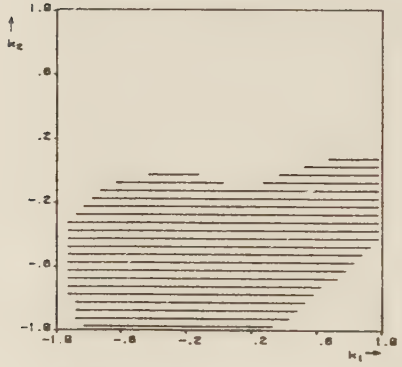


Bild A7.1b
Stabilitätsbereich
(Kein 4/8883-Mode) k_3= .2
APPDM(4)/(2/2/1.28) k_4= -.2

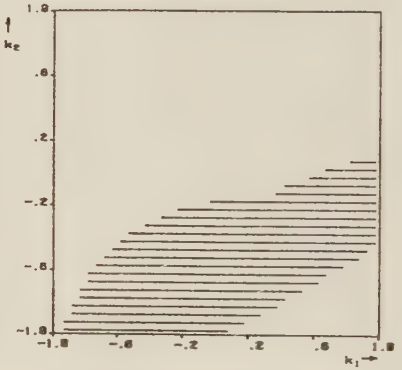


Bild A7.1e
Stabilitätsbereich
(Kein 4/8883-Mode) k_3= .4
APPDM(4)/(2/2/1.28) k_4= -.4

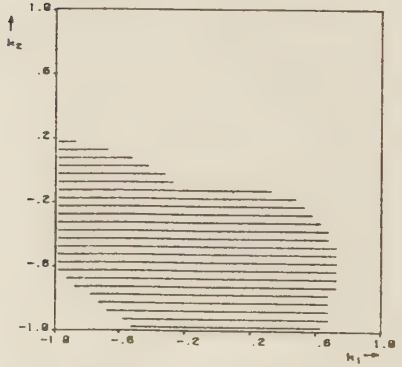


Bild A7.1c
Stabilitätsbereich
(Kein 4/8883-Mode) k_3= -.2
APPDM(4)/(2/2/1.28) k_4= -.2

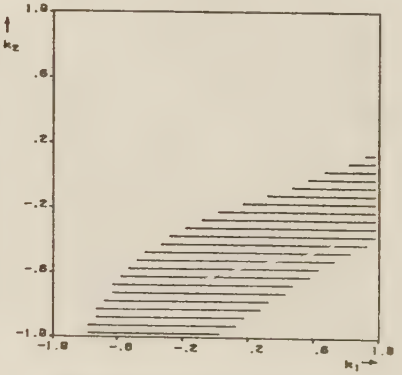


Bild A7.1f
Stabilitätsbereich
(Kein 4/8883-Mode) k_3= .5
APPDM(4)/(2/2/1.28) k_4= -.5

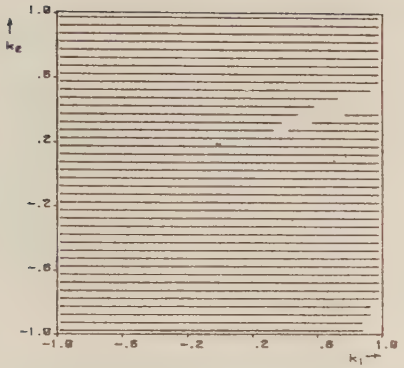


Bild A7.2a
Stabilitätsbereich
(Kein 6/8813-Mode) k_3= .2
APPDM(4)/(2/2/1.28) k_4= .2

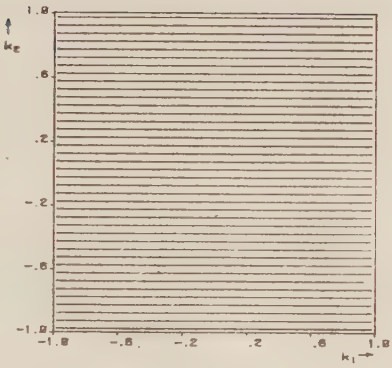


Bild A7.2d
Stabilitätsbereich
(Kein 6/8813-Mode) k_3= .3
APPDM(4)/(2/2/1.28) k_4= -.3

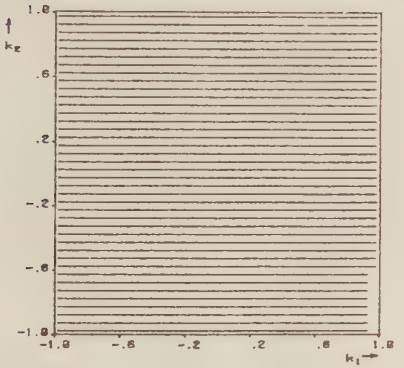


Bild A7.2b
Stabilitätsbereich
(Kein 6/8813-Mode) k_3= .2
APPDM(4)/(2/2/1.28) k_4= -.2

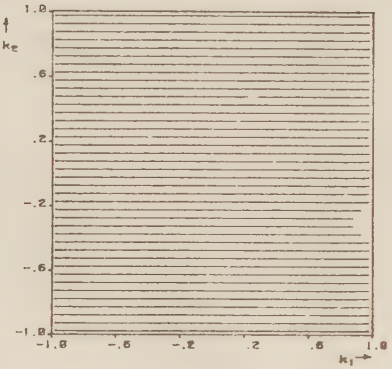


Bild A7.2e
Stabilitätsbereich
(Kein 6/8813-Mode) k_3= .4
APPDM(4)/(2/2/1.28) k_4= -.4

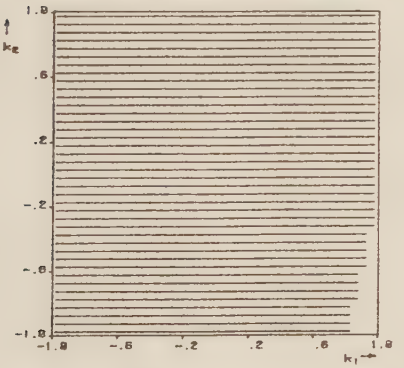


Bild A7.2c
Stabilitätsbereich
(Kein 6/8813-Mode) k_3= -.2
APPDM(4)/(2/2/1.28) k_4= -.2

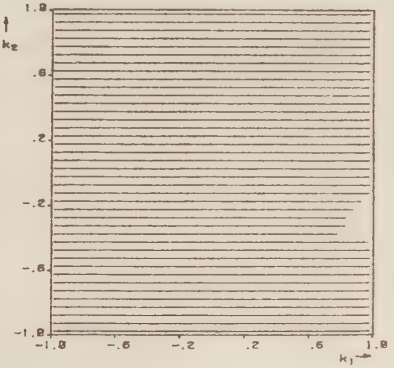


Bild A7.2f
Stabilitätsbereich
(Kein 6/8813-Mode) k_3= .5
APPDM(4)/(2/2/1.28) k_4= -.5

A8 Stabilitätsdiagramme zur Beurteilung der Existenzfähigkeit eingeprägter Moden

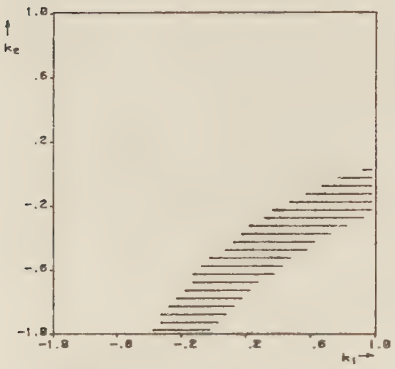


Bild A8.1a
Stabilitätsbereich
(Koln 4/0003-Mode) k_3= .885468
APPDM(4)/(2/2/1.17) k_4= -.906847

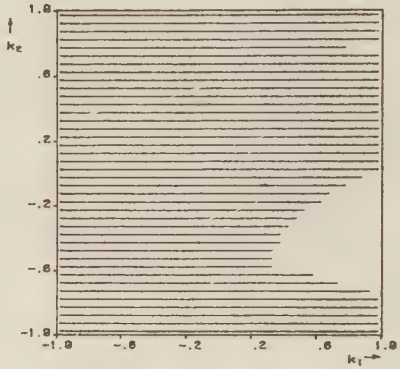


Bild A8.2a
Stabilitätsbereich
(Koln 5/0013-Mode) k_3= .973947
APPDM(4)/(2/2/1.25) k_4= -.984376

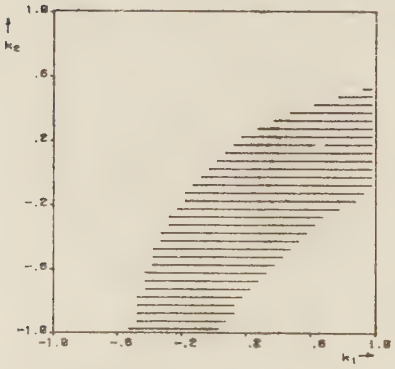


Bild A8.1b
Stabilitätsbereich
(Koln 4/0003-Mode) k_3= .986759
APPDM(4)/(2/2/1.17) k_4= .187588

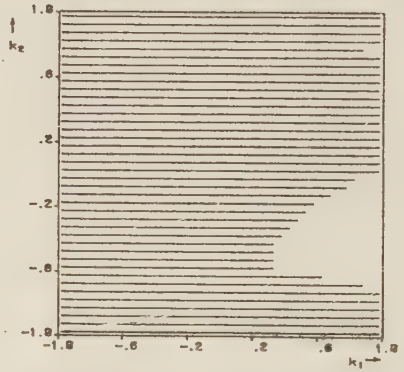


Bild A8.2b
Stabilitätsbereich
(Koln 5/0013-Mode) k_3= .986259
APPDM(4)/(2/2/1.25) k_4= -.988588

Lebenslauf

5. 7. 1945	geboren in Schleswig
1952 - 1956	Besuch der Volksschule in Essen-Steele
1956 - 1962	Besuch des Carl-Humann Gymnasiums in Essen-Steele und Abschluß mit der mittleren Reife
1962 - 1965	Ausbildung zum Elektromechaniker bei der Siemens AG in Essen
1965 - 1966	Tätigkeit als Elektromechaniker bei der Siemens AG in Essen
1966 - 1970	Studium der Elektrotechnik an der staatlichen Ingenieurschule in Duisburg. Während dieser Zeit (Juli und August 1969) Tätigkeit als Werkstudent bei der Siemens AG in München
26. 1. 1970 - 31. 3. 1970	Vorübergehende Tätigkeit als techn. Angestellter bei der Siemens AG in Essen
1970 - 1975	Studium der Elektrotechnik an der Ruhr-Universität Bochum
7. 2. 1975	Diplom-Hauptprüfung im Fach Elektrotechnik
seit 1. 4. 1975	Wissenschaftlicher Angestellter am Lehrstuhl für Nachrichtentechnik der Ruhr-Universität Bochum

DER DRUCKER - (Alte) Buscheystraße 138 - 4630 Bochum-Querenburg